STRUCTURE-FUNCTION STUDIES OF SAGEBRUSH FARNESYL DIPHOSPHATE SYNTHASE AND CHRYSANTHEMYL DIPHOSPHATE SYNTHASE BY CHIMERAGENESIS

by

Joseph Scott Lee

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

The University of Utah

August 2015

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Joseph Scott Lee**

has been approved by the following supervisory committee members:

C. Dale Poulter	, Chair	**March 5, 2015** Date Approved
Cynthia J. Burrows	, Member	**March 5, 2015** Date Approved
David P. Goldenberg	, Member	**March 5, 2015** Date Approved
Joel M. Harris	, Member	**March 5, 2015** Date Approved
Jennifer Heemstra	, Member	**March 5, 2015** Date Approved

and by **Cynthia J. Burrows**, Chair/Dean of

the Department/College/School of **Chemistry**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Despite differing catalytic specificities and activities, farnesyl diphosphate synthase (FPPase) and chrysanthemyl diphosphate synthase (CPPase) from *Artemisia tridentata* ssp. spiciformis have sequence alignments showing 69% identity and 84% similarity. The active sites of the enzymes are formed by a six-membered, α-helical bundle representative of the type-I isoprenoid synthase fold (IS-1 fold). FPPase is selective for the 1'-4 coupling (chain elongation) of dimethylallyl diphosphate (DMAPP) and two isopentenyl diphosphate (IPP) molecules, initially forming geranyl diphosphate (GPP) and then farnesyl diphosphate (FPP). CPPase preferentially forms GPP when chain elongating, and can also couple two DMAPP molecules (irregular coupling) to preferentially form the c1'-1-2 product chysanthemyl diphosphate (CPP). CPPase additionally produces the 1'-2 product lavandulyl diphosphate (LPP) and a trace amount of the c1'-2-3-2' product maconelliyl diphosphate (MPP). The catalytic diversity of CPPase comes at the cost of catalytic efficiency (k_{cat}/K_m), as FPPase chain elongates at over 30,000-fold greater efficiency. In this study chimeric enzyme constructs were built from the IS-1 folds of FPPase and CPPase. Each enzyme was turned into the other through sequentially swapping the helices and loops of their IS-1 fold, building enzymes of varying FPPase and CPPase character to assess what structural elements affect catalytic specificity and activity. The first catalytic transformation observed along the N- to C-terminal conversion of FPPase to CPPase was a shift from preferential FPP to GPP

formation. The GPPase showed over 2000-fold greater catalytic efficiency toward terminating chain elongation at the C_{10} product. Then, catalytic efficiency dropped to CPPase-like levels, correlating with a T194G FPPase to CPPase mutation of the KT motif in the fourth conserved region among E-chain elongation enzymes. Following, irregular terpenoid catalysis was observed in the form of preferential LPP formation, associated with F231Y and D235N mutations in the fifth conserved region. Preferential CPP production was dependent upon an enzyme having C-terminal sequence outside the IS-1 fold from CPPase. Replacement of the N-terminal region outside of the IS-1 fold of CPPase with FPPase sequence reclaimed GPP binding ability and FPP formation. A return to an FPPase-like catalytic efficiency was not observed in any chimera along the N- to C-terminal metamorphosis of CPPase to FPPase, further indicating the significance of the C-terminal region in the catalytic activity and specificity of the enzymes.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

aa	amino acid
Amp	ampicillin
ATP	adenosine triphosphate
β-Me	β-mercaptoethanol
bp	base pair(s)
BSA	bovine serum albumin
C^+	chrysanthemyl (cyclopropyl) cationic intermediate
CE^{DMAPP}	catalytic efficiency with respect to dimethylallyl diphosphate
CE^{GPP}	catalytic efficiency with respect to geranyl diphosphate
CE^{IPP}	catalytic efficiency with respect to isopentenyl diphosphate
CIP	calf intestinal phosphatase
CPP	chrysanthemyl diphosphate
CPPase	chrysanthemyl diphosphate synthase
DI	deionized
DMA	dimethyallyl
DMA^+	dimethylallyl cationic intermediate
DMAP	dimethyallyl phosphate
DMAPP	dimethyallyl diphosphate
$DMAPP^{EA}$	dimethylallyl diphosphate electron acceptor

$DMAPP^{ED}$	dimethyallyl diphosphate electron donor
DMASPP	thiodimethylallyl diphosphate
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide
DXP	1-deoxy-D-xylulose-5-phosphate
EA	electron acceptor
ED	electron donor
EDTA	ethylenediaminetetraacetic acid
E,E-F-OH	*E,E*-farnesol
E,E-FPP	*E,E*-farnesyl diphosphate
F-OH	farnesol
FPLC	fast protein liquid chromatography
FPP	farneyl diphosphate
FPPase	farnesyl diphosphate synthase
GC-FID	gas chromatography with flame ionization detection
G-OH	geraniol
GPP	geranyl diphosphate
GGPP	geranylgeranyl diphosphate
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HMB-PP	(E)-4-Hydroxy-3-methyl-but-2-enyl pyrophosphate
HMG-COA	3-hydroxy-3-methylglutaryl-CoA
IDI	isopentenyl diphosphate isomerase
iG-OH	3-methylene geraniol

iGPP	3-methylene geranyl diphosphate
IP	isopentenyl phosphate
IPK	isopentenyl phosphate kinase
IPP	isopentenyl diphosphate
IPTG	isopropyl β-D-thiogalactoside
IS-1	isoprenoid synthase type 1
L^+	lavandulyl (branched) cationic intermediate
LB	lysogeny broth
L-OH	lavandulol
LPP	lavandulyl diphosphate
MEGAWHOP	megaprimer whole plasmid
MEP	2-*C*-methyl-D-erythritol 4-phosphate
MEV	mevalonate
M-OH	maconelliol
MOPS	3-(N-morpholino)propanesulfonic acid
MPP	maconelliyl diphosphate
mRNA	messenger ribonucleic acid
MTBE	methyl tert-butyl ether
MWCO	molecular weight cut-off
NEB	New England Biolabs
N-OH	nerol
NPP	neryl diphosphate
OD_{550}	optical density at 550 nm

OD_{600}	optical density at 600 nm
PCR	polymerase chain reaction
P-OH	planococcol
PPP	planococcyl diphosphate
rt	room temperature
SCOPE	structure-based combinatorial protein engineering
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SOC	super optimal with catabolite repression
TAE	Tris-acetate-EDTA
TLC	thin layer chromatography
UV	ultraviolet
Z,E-F-OH	*Z,E*-farnesol
Z,E-FPP	*Z,E*-farnesyl diphosphate

ACKNOWLEDGEMENTS

I would like to sincerely thank Professor Dale Poulter for the opportunity to work and develop in his lab – I have personally and professionally learned much in my time here. And I would like to thank my wonderful wife, Whitney Irvine, for her love, support, and encouragement.

CHAPTER 1

INTRODUCTION

1.1 Charles Darwin and Alfred Wallace Had It Right

In 1869 the great naturalist and author John Muir penned "when we try to pick out anything by itself, we find it hitched to everything else in the Universe" in his journal.[1] Perhaps more truth and application lies in the thought than he realized. Earth touts an abundance of life, displaying it in bountiful forms and functions. Despite the apparently autonomous differences, the similarities arguably hold more significance – life utilizes the same heredity material and genetic code, with proteins built using that code and common amino acids. It is through the synonymy that nature exists. No matter how unique the appearance and functionality a creature appears, it finds itself hitched to all other life through fundamental biochemical processes.

In 1858 the idea of the evolution of species from a common ancestor at the hand of natural selection was copresented by Charles Darwin and Alfred Wallace. The action coined by Darwin as natural selection is, in short, the nature of organic beings to "multiply, vary, let the strongest live and the weakest die."[2] The rationale behind the selective action is that favorable variations are preserved in future generations due to an advantage in environmental utilization and proliferation. Natural selection was a hard sell for many scientists and struggled to find acceptance. It was not until the discovery of

genes and mutations in the 20th century that natural selection was not only alluring as an explanatory tool, but also required. We now know natural selection as part of divergent evolution, which produces new biological traits and nature's diversity through the duplication and modification, and domain swapping of genes.[3,4]

1.2 A Comment on the Interconnectivity of the Domains of Life

> The affinities of all the beings of the same class have sometimes been represented by a great tree. As buds give rise by growth to fresh buds, and these if vigorous, branch out and overtop on all sides many a feebler branch, so by generation I believe it has been with the great Tree of Life, which fills with its dead and broken branches the crust of the earth, and covers the surface with its ever branching and beautiful ramifications.
>
> Charles Darwin, On the Origin of Species, 1859

The earliest known life existed on Earth more than 3.7 billion years ago.[5] In that time a vast range of relatively simple to complex organisms evolved and inhabit this domain from one "unearthly" extreme to another. A phylogenetic tree is used to characterize the evolutionary relationships among life's creatures, with each branch point representing where divergence from a common ancestor occurred and closely related, yet distinct organisms continued on unique paths. Life is categorized into one of three Domains – Archaea, Bacteria, and Eukaryota (Figure 1.1).[6] The fundamental difference among organisms is the presence or lack of membrane-bound organelles. Archaea and Bacteria lack organelle membranes and are termed prokaryotic, whereas eukaryotic organisms have, at minimum, their genetic material encapsulated in a membrane structure we know as the nucleus.

The Archaea and Bacteria domains consist of microorganisms typically a few micrometers in length and varying in shape. Archaebacteria are commonly thought of as extremophiles, living in hot, salty, or acidic environments too inhospitable for most

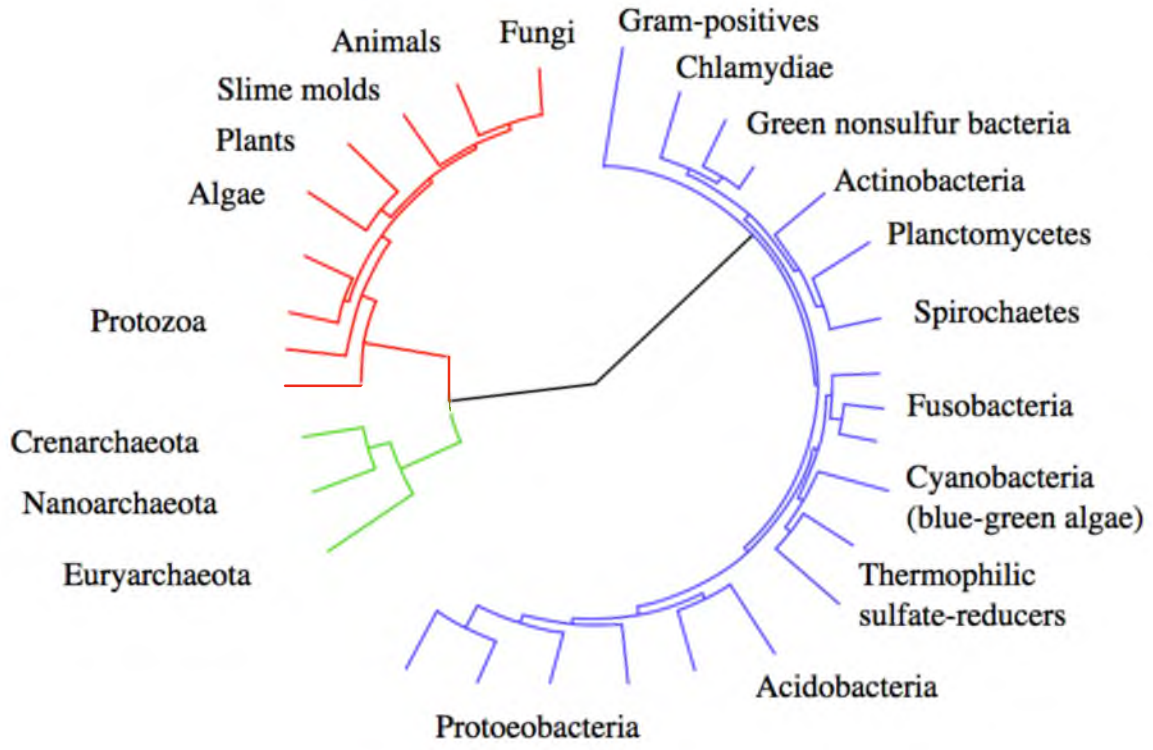


Figure 1.1. A highly resolved Tree of Life, derived from completely sequenced genomes, distinguishing between the three domains of life. Eukaryotes are colored red, bacteria blue, and archaea green.[7]

beings, yet are found in a wide range of habitats, including the human navel.[8] Extremophilic archaea, especially those resistant to heat or pH extremes, are a source of enzymes capable of tolerating the harsh conditions, for example, the revolutionizing thermostable DNA polymerases used for quick and easy DNA cloning. Akin to their archaeal counterparts, bacteria are found in nearly every corner of the earth. In fact, there are roughly 40 million bacterial cells in a gram of soil, and one million in a milliliter of water.[9] Surprisingly, the biomass of bacteria on Earth outweighs that of plants and animals combined.[10] The versatility of bacteria are harnessed in a variety of industrial processes such as cheese and yogurt production, sewage treatment, and antibiotic production. The adaptability of bacteria is unmatched, and therefore their ability to thrive

in a multitude of environments.

Although there are numerous unicellular eukaryotic organisms, such as the protozoa, we tend to be more aware of our multicellular eukaryotic counterparts – mainly those from the kingdoms Plantae, Animalia, and Fungi. This awareness is not without good reason, a rich evolutionary history exists among these kingdoms. Long have animals relied on other animals, plants, and fungi for sustenance, and vice versa. In that, it is sometimes easy to lose perspective of the entire picture, there is much more than meets the eye. For example, humans have ten times more microbial cells than their own, and that microbiome affects our physical and psychological well-being.[11-13] Humans may be more aptly described as superorganisms whose metabolism is a combination of microbial and human facets, in fact, it appears most organisms are just one part of an interdependent metaorganism.[14,15] Since the beginning of life some 3.7 billion years ago organisms have evolved and diversified together, formed symbioses and battled competitively. Over the years, interactions of give and take have molded the species into what lives and exists today, each with a uniqueness celebrating the long evolutionary struggle to multiply, diversify, and thrive in ever-changing environmental niches.

1.3 The Exceptional Biosynthesis of Secondary Metabolites by Plants

Plants are specialized in their ability to photosynthesize through harvesting light energy, carbon dioxide, and water to build chemical energy in the form of sugar, and in doing so release oxygen. Through the maintenance of atmospheric oxygen levels and supply of energetic organic compounds, photosynthesis is arguably the most important process to supporting life on Earth.[16] Plants exhibit an exceptional ability to synthesize

structurally and chemically diverse low molecular weight compounds, numbering into the hundreds of thousands.[17] Few of the compounds are deemed as "primary" metabolites, which are those produced ubiquitously by all plant species and other organisms (sugars, amino acids, nucleotides, lipids, and energy sources).[18] The remaining "secondary" metabolites are described as compounds not directly involved in the growth, development, or reproduction of an organisms.

It is generally agreed that phytochemistry began when Friedrich Wilhelm Serturner isolated morphine from *Papaver somniferum* (commonly known as the Opium poppy) over 200 years ago (Figure 1.2).[17] The attribution of the active component in a plant drug to a single chemical compound began natural (secondary) product chemistry. One active secondary metabolite isolate after another was identified, and swift progress in natural product research fueled growth in synthetic, analytical, and pharmaceutical chemistry. The synthesis of indigo, a plant secondary metabolite, by Adolf Baeyer and Viggo Drewson in 1882 marks a milestone in synthetic organic chemistry. In continuation with the classic opioid, Robert Robinson elucidated morphine's structural formula in the 1920s, with the total synthesis completed by Marshall Gates in 1956.[19] The ability to map plant biosynthetic pathways using radiolabeled tracer techniques began in the mid-20th century with the introduction of ^{14}C and ^{3}H labeled compounds, and in the following decades the pathways for many of the significant secondary metabolite classes were defined.[17]

Secondary metabolites were selected over the course of evolution to address specific needs, many due to the sessile nature of plants. The colors and scents of flowers are evolved to attract insect pollinators to increase the chance of fertilization.[20,21] Toxic

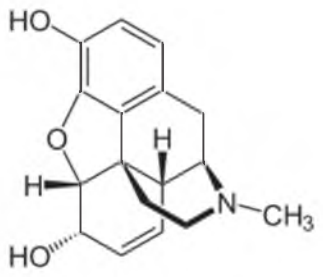


Figure 1.2. Structure of morphine.

chemicals provide defense from pathogens and herbivores, or may inhibit the growth of neighboring plants.[22-25] Chemicals in fruit provide delightful scents, colors, and flavors to promote animal consumption and seed dispersal. Furthermore, humankind has utilized natural products for thousands of years for a wide range of purposes including flavors, fragrances, dyes, stimulants, insecticides, hallucinogens, poisons, and therapeutics.[18]

The chemical solutions to common obstacles vary among plant lineages.[4] The scents of flowers greatly differ from species to species, and although the compounds synthesized to deter herbivores are vast, individual lineages only produce a small portion of them.[26,27] Plant genomes contain an estimated 20,000 – 60,000 genes, with 15-25% of those genes encoding enzymes for secondary metabolism.[4]

Natural products can be divided into three main classes – phenolic compounds, terpenoids/isoprenoids, and nitrogen or sulfur containing compounds (such as alkaloids and glucosinolates).[18] As expected from an evolutionary standpoint, the three main classes of secondary metabolites are biosynthetically linked to primary metabolic pathways (Figure 1.3). Genes for secondary metabolites may evolve from already existing natural product genes, or genes for primary metabolites; nonetheless, it is likely the ancestor gene was associated with primary metabolism, which serves as a pool for

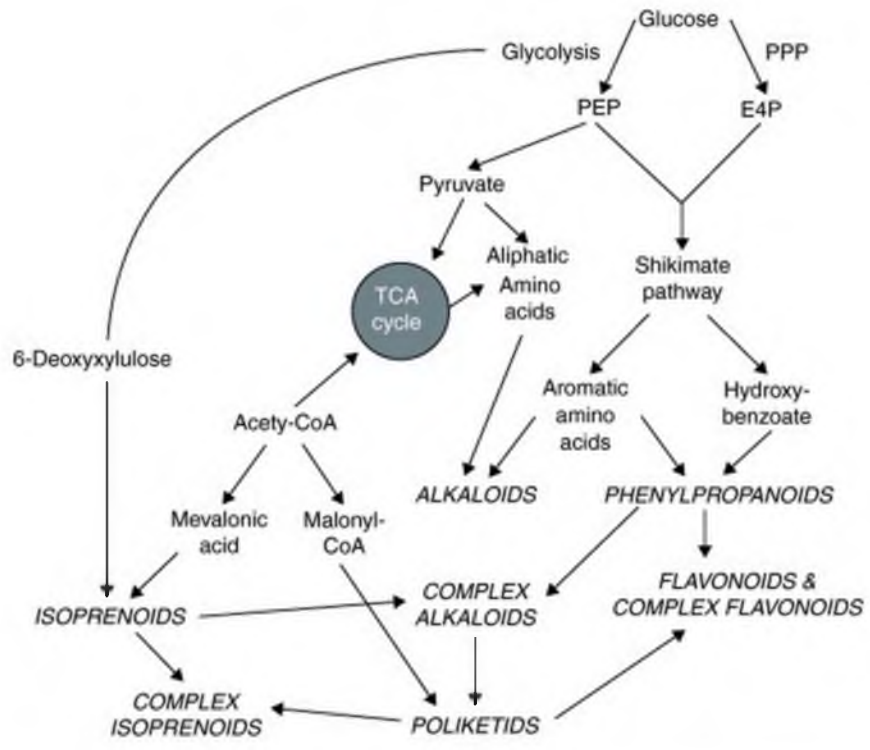


Figure 1.3. The biosynthesis of plant secondary metabolites is closely linked to primary metabolic pathways.[18]

new secondary metabolite-related genes.[4,18]

1.4 The *E*-Terpenoid Biosynthetic Pathway

The isoprenoid biosynthetic pathway is one of nature's largest and most diverse. With over 62,000 natural products, terpenes include many well-known groups including sterols, ubiquinones, carotenoids, and essential oils.[28-30] Despite responsibility for a wide structural array of chemicals, isoprenoids are simply built from coupling 3-methyl-1-butyl (isoprene) units. Dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP) are the universal isoprene units coupled to make terpenoid skeletal backbones in one of four ways – chain elongation, branching, cyclopropanation, and cyclobutanation.[29]

Chain elongation is considered a regular or head-to-tail coupling, while the others are irregular or nonhead-to-tail (Figure 1.4). Chain elongation is by far the most common coupling reaction, showing ubiquity across organisms. Two families of regular chain elongation enzymes exist and are grouped based upon the stereochemistry of the double bond in the added isoprene unit – *E* or *Z*. Generally, *E*-polyprenyl diphosphate synthases produce short-chain products early in the biosynthetic pathway, whereas *Z*-family enzymes make longer-chain products, notably the 10^5-10^6 Dalton compound natural rubber.[31]

If the *E*-terpenoid pathway is viewed as a tree, the mevalonate (MEV) and non-mevalonate (or methylerythritol/deoxyxylulose phosphate, MEP/DXP) pathways are the roots (Figure 1.5). These independent, nonhomologous pathways produce nature's source of DMAPP and IPP. Most eukaryotes, archaea, and some eubacteria utilize the MEV pathway.[32] Three acetyl-Coenzyme A molecules are condensed and converted into IPP, from which the enzyme isopentenyl diphosphate isomerase (IDI) generates DMAPP. The MEV pathway is highly relevant to human health – inhibiting 3-hydroxy-3-methyl-glutaryl-CoA (HMG-CoA) reductase by statins limits sterol production, thereby lowering cholesterol and blood pressure, among other potential health benefits.[33] Until relatively recently, the MEV pathway was thought of as the straightforward, sole source of the foundational isoprene units, but continued discoveries indicate the MEV pathway is more complex than originally thought.[34-36] The MEP pathway appears in cyanobacteria, algae, eubacteria, plant chloroplasts, and apicomplexan parasites.[32] Pyruvate and glyceraldehyde-3-phosphate are condensed and converted into the terpenoid building blocks. The MEP pathway does not require IDI as DMAPP and IPP are both produced

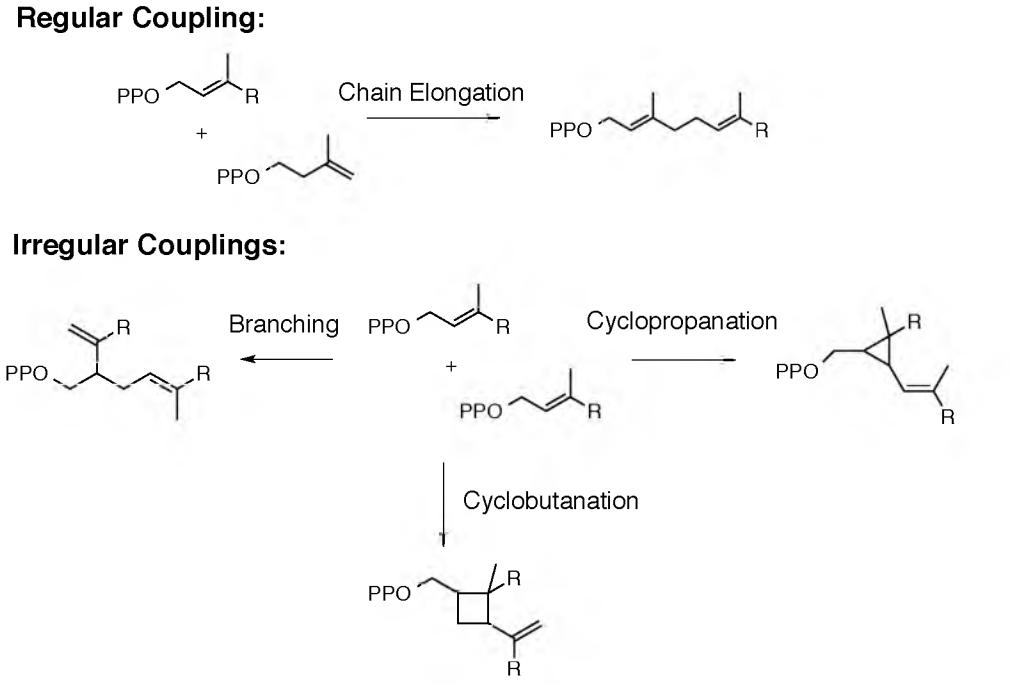


Figure 1.4. Isoprenoids are built from four fundamental coupling reactions – chain elongation, branching, cyclopropanation, and cyclobutanation.

from the final substrate via (*E*)-4-hydroxy-3-methyl-but-2-enyl pyrophosphate (HMB-PP) reductase.

The trunk of the pathway, which symbolizes chain elongation, extends from the tree's roots. In the *E*-chain elongation pathway, the first coupling sequence linearly links DMAPP and IPP (1'-4, respectively) to produce the monoterpene (C_{10}) geranyl diphosphate (GPP). The prenyl transfer reaction occurs through a dissociative electrophilic alkylation, finalized by extraction of the pro-*R* proton at C2 of the IPP unit to produce the *E* stereochemistry (Figure 1.6).[37,38] GPP can linearly couple with another IPP molecule and extend to the sesquiterpene (C_{15}) farnesyl diphosphate (FPP), and so on up the trunk. This ability to sequentially extend the chain with IPP is central to the isoprenoid pathway. In addition to providing biological functions, for example, the

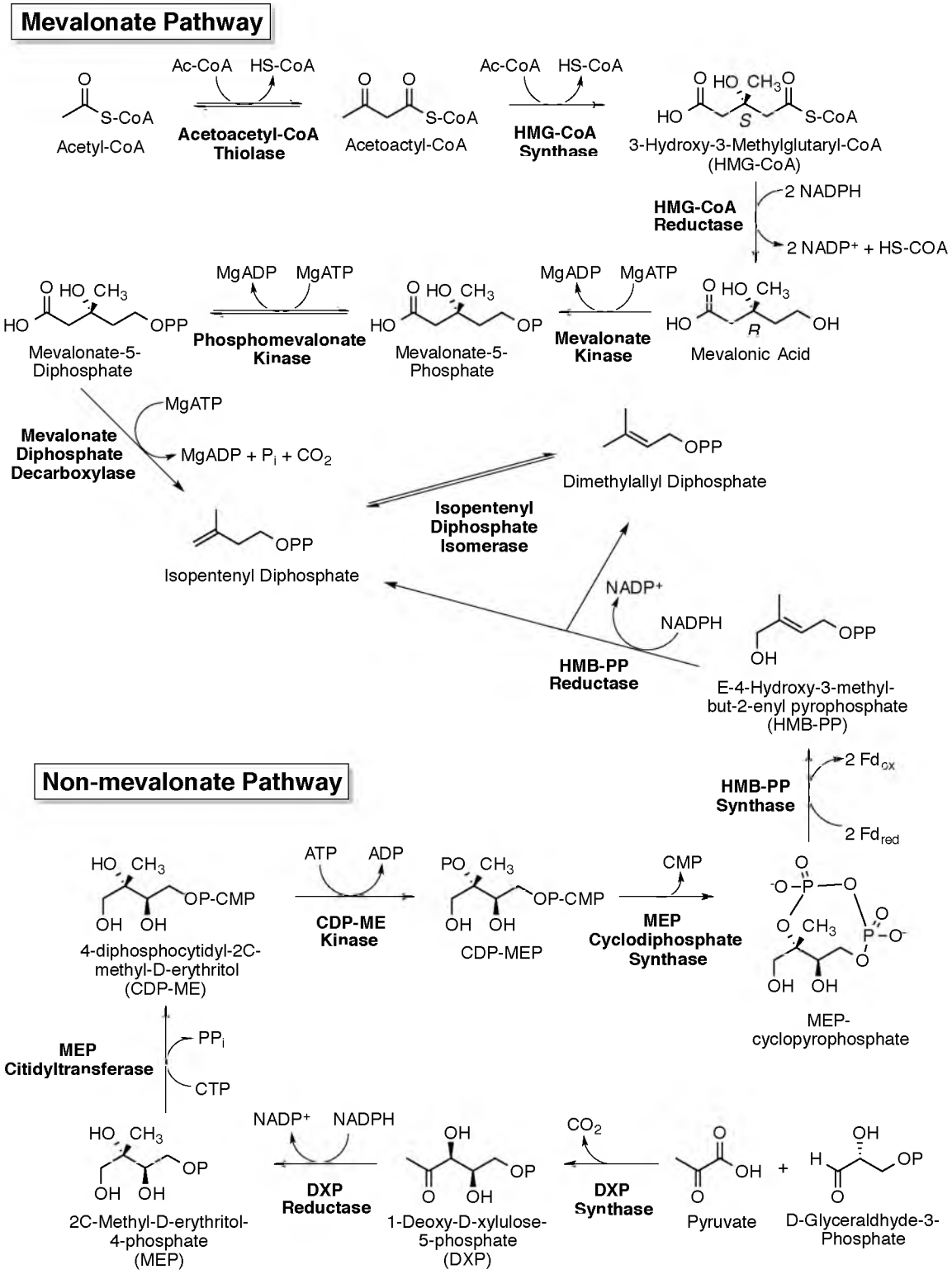


Figure 1.5. The biosynthesis of DMAPP and IPP through the mevalonate and non-mevalonate pathways.

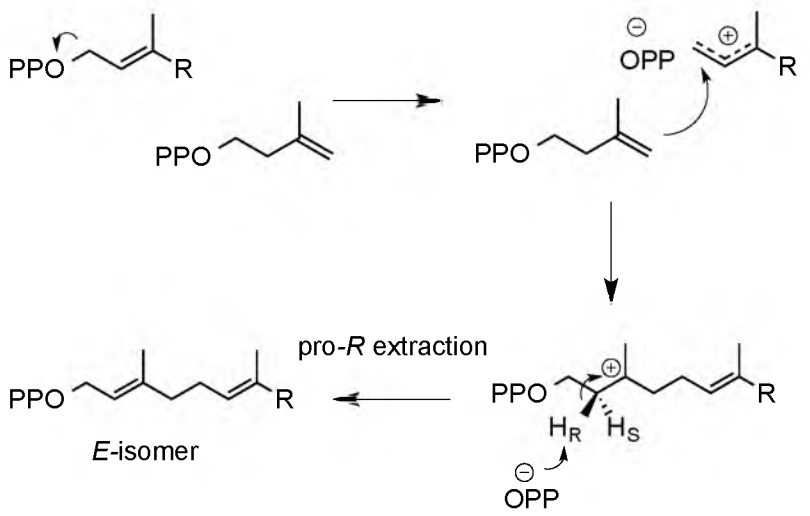


Figure 1.6. The chain elongation reaction occurs through a dissociative electrophilic alkylation.

prenylation of proteins by FPP or GGPP for protein-membrane anchoring and cell signal transduction, each chain elongation product is a branching point in the biological pathway.[39] A linear isoprenoid can cyclize or irregularly couple with itself to produce an array of structures that may undergo further modification and specification, branching out from the tree's trunk and creating an isopenoid subclass, such as the sterols, carotenoids or gibberellins (branching from FPP or GGPP, respectively) (Figure 1.7).

Nature exhibits several isoprenoid backbone structures outside of the four fundamental couplings (Figure 1.8). Some are found in limited cases, the 4'-4 (tail-to-tail) coupling is present in archaeal cell membranes composed of the diterpene phytane and the 1'-2 (branched) coupling is exhibited in select plants.[40-42] Also, the c1'-2-3-2' and c2'-2-3-4-3' connectivities (cyclobutyl and cyclopentyl) show near exclusive production as mealybug pheromones.[43-45] Most irregular couplings traverse through a c1'-2-3 (cyclopropyl) intermediate via a dissociative electrophilic alkylation mechanism. The finding that presqualene diphosphate, a cyclopropyl triterpene, was an intermediate in the

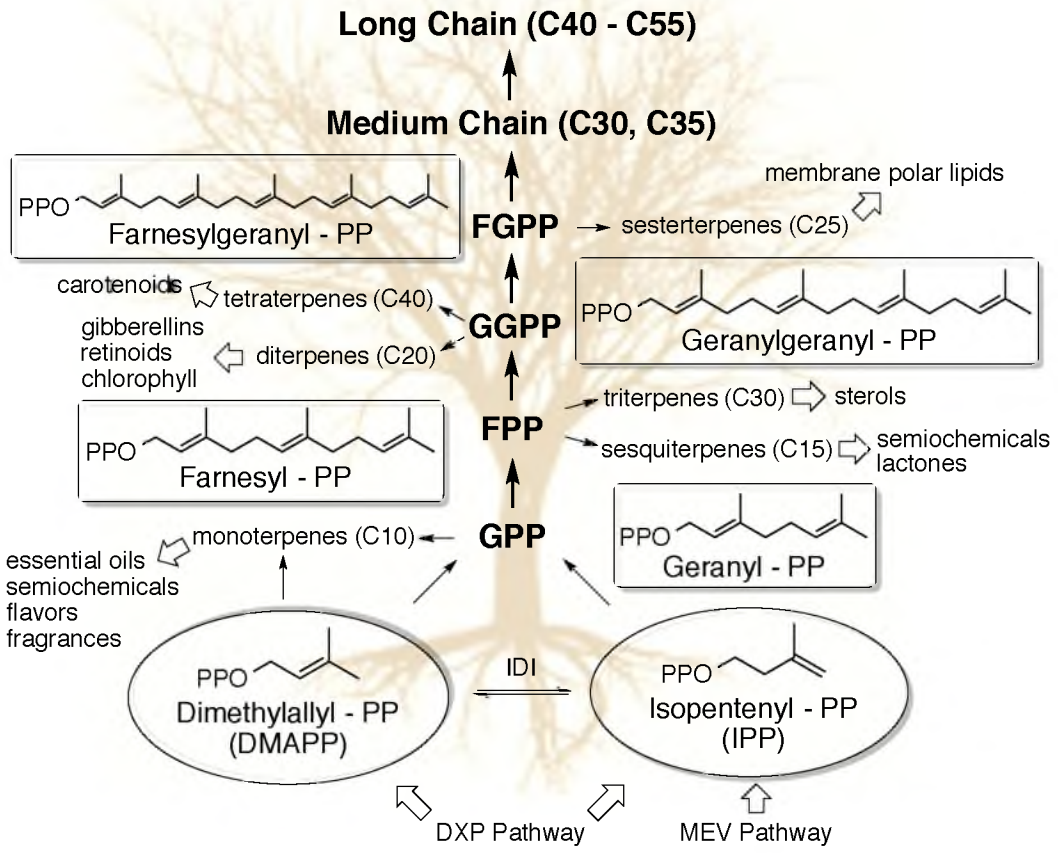


Figure 1.7. Overview of the *E*-chain elongation pathway used for the biosynthesis of terpenoids (DXP, deoxyxylulose; MEV, mevalonate; IDI, isopentenyl diphosphate isomerase).

biosynthesis of the 1'-1 linked squalene initiated great thought on the biogenesis of non-head-to-tail couplings.[46,47] Soon after, it was discovered that the cyclopropyl-coupled prephytoene, a C_{40} analogue of presqualene, was the intermediate between geranylgeranyl diphosphate and the 1'-1 coupled product phytoene; and it was recognized and demonstrated that c1'-2-3 cationic rearrangements led to the formation of 1'-3, 2-1'-3, and c1'-1-2 couplings.[48-50] Then strong evidence for a cyclopropyl intermediate in the biosynthesis of the 1'-3 coupled triterpene botryococcene was

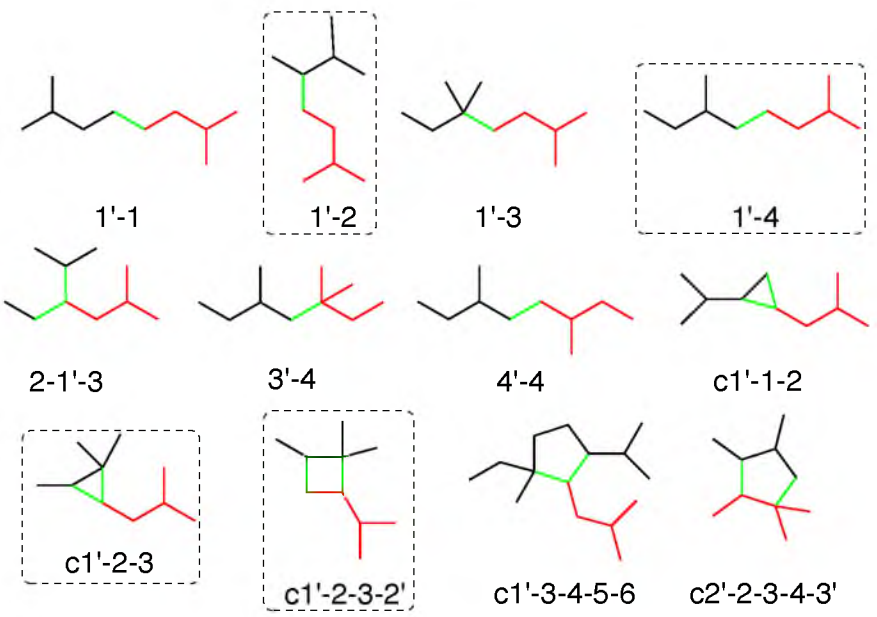


Figure 1.8. Isoprenoid coupling patterns observed in nature. The four fundamental couplings are enclosed in dashed lines.

obtained through stable-isotope labeling experiments.[51,52] Eventually a chrysanthemyl diphosphate synthase was isolated that demonstrated the ability to produce all four fundamental coupled products, and does so through the cyclopropyl intermediate (Figure 1.9).[29] Finally, botryococcene-like compounds reported from organic-rich lakebed sediments and plant extracts with a peculiar cyclopentane coupling (C1'-3-4-5-6) was recently identified as a promiscuous product from a recombinant squalene synthase. [53-55]

One could argue that chain elongation and cyclopropanation are the sole fundamental couplings. Chain elongation is undoubtedly a fundamental coupling due to its uniqueness and pivotal role in the isoprenoid pathway, and cyclopropanation due to its formation and rearrangement as a prerequisite for most other couplings. Regardless, the complexity and diversity of products originating from the union of relatively simple five carbon units is not only impressive, but also utilized by all forms of life except a small

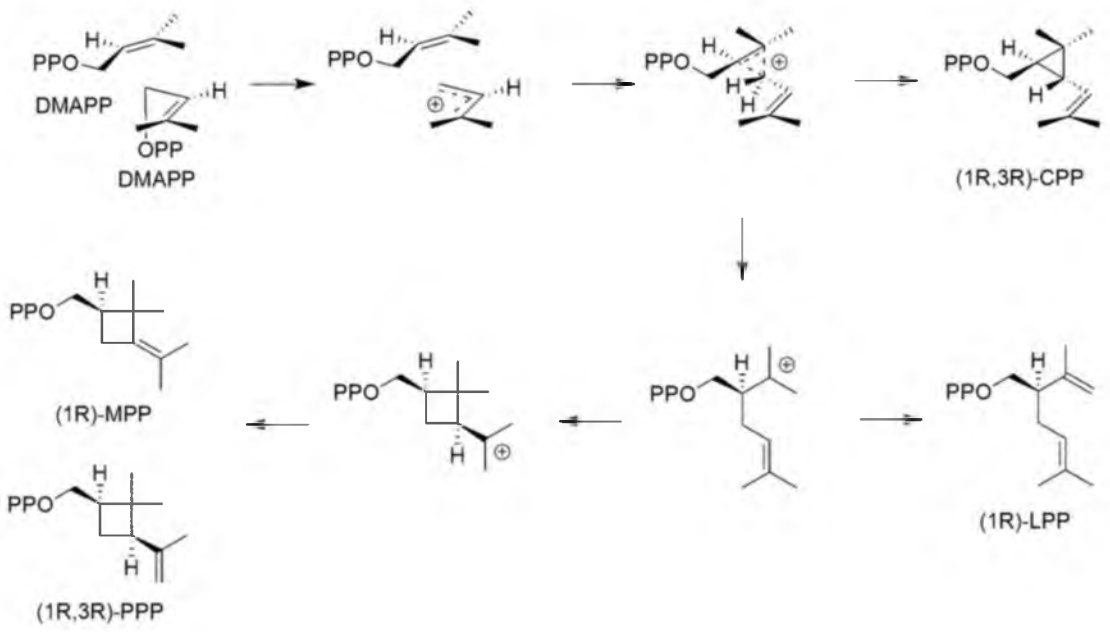


Figure 1.9. The fundamental, irregularly coupled products of the isoprenoid biosynthetic pathway utilize a common mechanism that initially traverses through a cyclopropyl, carbocationinc intermediate (DMAPP, dimethyallyl diphosphate; CPP, chrysanthemyl diphosphate; LPP, lavandulyl diphosphate; MPP, maconelliyl diphosphate; PPP, planococcyl diphosphate).

group of parasitic bacteria.[57] Nature's development and evolution of the terpenoid biosynthetic pathway provides a marvelous and rich source of industrially and pharmacologically relevant secondary metabolites, in addition to a seemingly endless base for scientific research.

1.5 The Type-I Isoprenoid Synthase Fold: Sage Farnesyl and Chrysanthemyl Diphosphate Synthases

The type I isoprenoid synthase (IS-1) fold is intimately involved in the *E*-chain elongation pathway. The all α-helical bundle and connective loops of the IS-1 fold were first observed in the x-ray structure of avian farnesyl diphosphate synthase (FPPase)

(Figure 1.10).[58] Since its debut in FPPase, the α-helical fold has been observed in other *E*-polyprenyl diphosphate synthases, enzymes catalyzing nonhead-to-tail reactions, and terpenoid cyclases, making enzymes exhibiting the structure members of a superfamily.[29,59-62] Currently over 16,000 IS-1 fold sequences reside in databases and the number continues to increase.[63] IS-1 enzymes utilize DDxxD/E ("aspartate rich") amino acid motifs to coordinate Mg^{2+} ions in a trinuclear complex that binds substrate diphosphates and catalyzes ionization.[62,64,65] Pyrophosphate departure triggers formation of carbocations, which the enzymes elegantly stabilize, protect while chaperoning rearrangement, and subsequently quench to finalize product formation.

The prevalence of the IS-1 fold suggests divergent evolution from a common primordial ancestor early in the evolution of terpene biosynthesis. Enzymes with such structural homology catalyzing subsequent steps in a biosynthetic pathway likely share an evolutionary origin.[66] Strong support for the idea was obtained with the isolation of closely related farnesyl and chrysanthemyl diphosphate synthases (FPPase and CPPase, respectively) from snowfield sagebrush *Artemisia tridentata* ssp. spiciformis, both enzymes being IS-1 fold homodimers.[67] FPPase is specific for the production of the regularly coupled C_{15} compound farnesyl diphosphate, whereas CPPase is catalytically promiscuous and exhibits all the fundamental couplings in its products (Figure 1.11). Despite an apparent catalytic gap between the enzymes, their sequence alignments show a striking 69% identity and 84% similarity, not only implying a recent divergence of the genes from a common origin, but also perhaps the evolution of cyclopropanation activity from an ancestral *E*-chain elongation enzyme (Figure 1.12).[56]

It is not surprising snowfield sagebrush evolved an enzyme with the synthetic

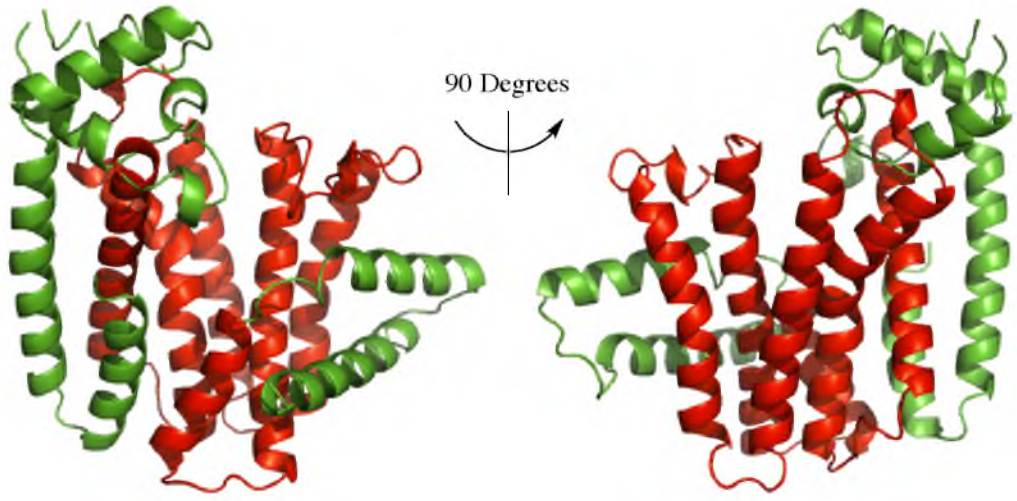


Figure 1.10. Monomeric crystal structure of sagebrush FPPase. Helices and loops of the IS-1 fold are colored in red.

capacity of CPPase, since closely related members of the plant group boast the greatest variety of nonhead-to-tail terpenes.[52] The distinguishing factor between the sagebrush enzymes is the ability of FPPase to exclusively bind IPP in the electron donor site for the prenyltransfer reaction. The *E*-polyprenyl chain elongases exhibit five conserved amino acid regions, and CPPase differs from FPPase by only two amino acids among the residues identified as important for substrate binding and activity in FPPase.[67,68] The staggering sequence (and therefore structural) resemblance, yet contrasting catalytic activities between the sagebrush enzymes yields a unique opportunity to probe structure-function aspects of the IS-1 fold.

1.6 The Exploration of Structure-Function Relationships of the IS-1 Fold Through *A. tridentata* FPPase and CPPase Chimeras

The use of chimeric proteins has proven utility in the enzymologist's toolbox. Chimeras can help deduce protein structure/function relationships, identify important

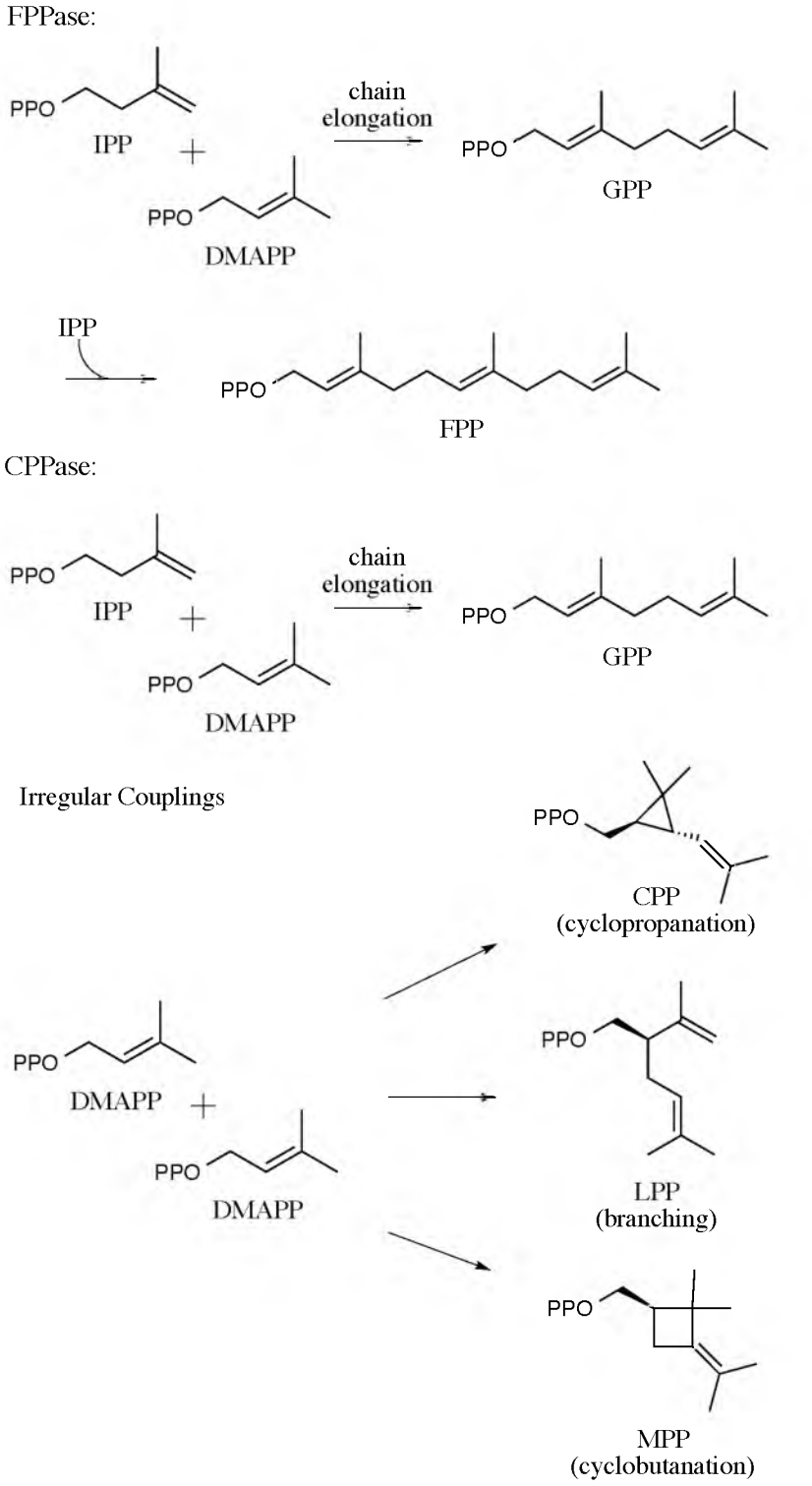


Figure 1.11. Reactions catalyzed by sagebrush FPPase and CPPase (IPP, isopentenyl diphosphate; DMAPP, dimethyallyl diphosphate; GPP, geranyl diphosphate; FPP, farnesyl diphosphate; CPP, chrysanthemyl diphosphate; LPP, lavandulyl diphosphate; MPP, maconelliyl diphosphate).

Figure 1.12. Sage FPPase and CPPase amino acid sequence alignments illustrating the helices forming the structure of the enzymes, specifically the six helices of the IS-1 fold. FPPase and CPPase amino acid sequence alignments show 69% identity and 84% similarity. Identical residues are in red, similar in blue, and different in black. The five conserved regions of E-prenyl chain elongation enzymes are highlighted in gray.

residues contributing to a desired function, and produce proteins for metabolic engineering.[69-71] In the case of FPPase and CPPase chimeragenesis, the tool is used to mimic evolution using DNA sequences that underwent the scrutiny of natural selection.

The first set of chimeras constructed from the sagebrush enzymes was based on substitutions of the five conserved *E*-polyprenyl chain elongation regions. Restriction sites flanking the five conserved regions were engineered into the sagebrush genes and chimeric proteins made via "cutting and pasting" the desired DNA sequence segments – FPPase was turned into CPPase piece by piece.[29] In parallel with the changes in sequence, the chimeras showed a transition in their product ratios. As FPPase transitioned into CPPase, GPP became the primary chain elongation product over FPP, and irregular

activity went from none, to cyclobutanation/branching, to solely branching, to branching/cyclopropanation.[55] Essentially, it appears the templates favoring catalysis of one fundamental coupling over another lie in the sequences of FPPase and CPPase.

The work presented in this dissertation is a continuation of sagebrush FPPase/CPPase chimeragenesis. This set of chimeras attempts to further deduce the structural elements in the IS-1 fold responsible for catalytic selectivity and activity. Instead of the five conserved *E*-chain elongation conserved regions, the chimeras were built around the α-helices and loops of the IS-1 fold. Chimeras were also assembled in both directions (turning FPPase into CPPase, and vice-versa) and done so using the polymerase chain reaction (PCR) in order to produce constructs that maintain wild-type enzymatic sequences. Product ratio and kinetic analyses were performed, if possible, in addition to molecular modeling of the parental enzymes and select chimeric enzyme constructs.

1.7 References

(1) Muir, J. *My First Summer in the Sierra;* Houghton Mifflin: Boston, 1911.

(2) Darwin, C.; Beer, G. *The Origin of Species;* Oxford University Press: Oxford, 1996.

(3) Doolittle, R. F. The Multiplicity of Domains in Proteins. *Annu. Rev. Biochem.* **1995**, *64*, 287-314.

(4) Pichersky, E.; Gang, D. R. Genetics and Biochemistry of Secondary Metabolites in Plants: An Evolutionary Perspective. *Trends Plant Sci.* **2000**, *5*, 439-445.

(5) Ohtomo, Y.; Kakegawa, T.; Ishida, A.; Nagase, T.; Rosing, M. T. Evidence for Biogenic Graphite in Early Archaean Isua Metasedimentary Rocks. *Nature Geosci.* **2014**, *7*, 25-28.

(6) Woese, C. R.; Kandler, O.; Wheelis, M. L. Towards a Natural System of Organisms: Proposal for the Domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acd. Sci. USA* **1990**, *87*, 4576-4579.

(7) Ciccarelli, F. D.; Doerks, T.; Mering, C. V.; Creevey, C. J.; Snel, B.; Bork, P. Toward Automatic Reconstruction of a Highly Resolved Tree of Life. *Science* **2006**, *311*, 1283-1287.

(8) Hulcr, J.; Latimer, A. M.; Henley, J. B.; Rountree, N. R.; Fierer, N.; Lucky, A.; Lowman, M.D.; Dunn, R. R. A Jungle in There: Bacteria in Belly Buttons are Highly Diverse, but Predictable. *PLoS ONE* [online] **2012**, *7*, Article Archive. http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0047712 (accessed 01/13/14).

(9) Whitman, W. B.; Coleman, D. C.; Wiebe, W. J. Prokaryotes: The Unseen Majority. *P. Natl. Acad. Sci. USA* **1998**, *95*, 6578-6583.

(10) Hogan, M. C. The Encyclopedia of Earth. http://www.eoearth.org/view/article/150368/ (accessed 04/06/14), under Biodiversity/Bacteria categories.

(11) Savage, D. C. Microbial Ecology of the Gastrointestinal Tract. *Ann. Rev. Microbiol.* **1977**, *31*, 107-133.

(12) Desbonnet, L.; Garrett, L.; Clarke, G.; Kiely, B.; Cryan, J. F.; Dinan, T. G. Effects of the Probiotic *Bifidobacterium infantis* in the Maternal Separation Model of Depression. *Neuroscience* **2010**, *170*, 1179-1188.

(13) Aas, J.; Gessert, C. E.; Bakken, J. S. Recurrent *Clostridium difficile* Colitis: Case Series Involving 18 Patients Treated With Donor Stool Administered Via a Nasogastric Tube. *Clin. Infect. Dis.* **2003**, *36*, 580-585.

(14) Gill, S. R.; Pop, M.; DeBoy, R. T.; Eckburg, P. B.; Turnbaugh, P. J.; Smauel, B. S.; Gordon, J. I.; Relman, D. A.; Fraser-Liggett, C. M.; Nelson, K. E. Metagenomic Analysis of the Human Distal Gut Microbiome. *Science* **2006**, *312*, 1355-1359.

(15) Bosch, T. C. G.; McFall-Ngai, M. J. Metaorganisms As The New Frontier. *Zoology* **2011**, *114*, 185-190.

(16) Bryant, D. A.; Frigaard, N-U. Prokaryotic Photosynthesis and Phototrophy Illuminated. *Trends Microbiol.* **2006**, *14*, 488-496.

(17) Hartmann, T. From Waste Products to Ecochemicals: Fifty Years Research of Plant Secondary Metabolism. *Phytochemistry,* **2007**, *68*, 2831-2846.

(18) Aharoni, A.; Galili, G. Metabolic Engineering of the Plant Primary-Secondary Metabolism Interface. *Curr. Opin. Biotech.* **2011**, *22*, 239-244.

(19) Gates, M.; Tschudi, G. The Synthesis of Morphine. *J. Am. Chem. Soc.* **1956**, *78*, 1380-1393.

(20) Dudareva, N.; Pichersky, E. Biochemical and Molecular Genetic Aspects of Floral Scent. *Plant Physiol.* **2000**, *122*, 627-633.

(21) Mol, J.; Grofewold, E.; Koes, R. How Genes Paint Flowers and Seeds. *Trends Plant Sci.* **1998**, *3*, 212-217.

(22) Bennet, R. N.; Wallsgrove, R. M.; Secondary Metabolism in Plant Defense – Mechanisms. *New Phytol.* **1994**, *127*, 617-633.

(23) Harborne, J. B. The Comparative Biochemistry of Phytoalexin Induction in Plants. *Biochem. Sys. Ecol.* **1999**, *27*, 335-367.

(24) Dixon, R. A,; Lamb, C. J.; Masoud, S.; Sewalt, V. J. H.; Paiva, N. L. Metabolic Engineering: Prospects for Crop Improvement through the Genetic Manipulation of Phenylpropanoid Biosynthesis and Defense Response – A Review. *Gene* **1996**, *179*, 69-71.

(25) Mitchell-Olds, T.; Gershenzon, J.; Baldwin, L.; Boland, W. Chemical Ecology in the Molecular Era. *Trends Plant Sci.* **1998**, *3*, 362-365.

(26) Knudsen, J. T.; Tollsten, L. Trends in Floral Scent Chemistry in Pollination Syndromes: Floral Scent Composition in Moth-Pollinated Taxa. *Bot. J. Linn. Soc.* **1993**, *113*, 263-284.

(27) Ananthakrishnan, T. N. Induced Responses, Signal Diversity and Plant Defense: Implications in Insect Phytophagy. *Curr. Sci.* **1999**, *76*, 285-290.

(28) Dictionary of Natural Products. http://dnp.chemnetbase.com/dictionary-search/results.do?id=7816252&props=&struct=start&disp=&si (accessed 01/08/14).

(29) Thulasiram, H. V.; Erickson, H. K.; Poulter, C. D. A Common Mechanism for Branching, Cyclopropanation, and Cyclobutanation Reactions in the Isoprenoid Biosynthestic Pathway. *J. Am. Chem. Soc.* **2008**, *130,* 1966-1971.

(30) Tarshis, L. C.; Proteau, P. J.; Kellogg, B. A.; Sacchettini, J. C.; Poulter, C. D. Regulation of Product Chain Length by Isoprenyl Diphosphate Synthases. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 15018-15023.

(31) Paterson-Jones, J. C.; Gilliland, M. G.; Van Staden, J. The Biosynthesis of Natural Rubber. *J. Plant Physiol.* **1990**, *136*, 257-263.

(32) Hunter, W. N. The Non-mevalonate Pathway of Isoprenoid Precursor Biosynthesis. *J. Biol. Chem.* **2007**, *282*, 21573-21577.

(33) Liao, J. K.; Laufs, U. Pleiotropic Effects of Statins. *Annu. Rev. Pharmacol.* **2005**, *45*, 89-118.

(34) Rohmer, M. The Discovery of a Mevalonate-independent Pathway for Isoprenoid Biosynthesis in Bacteria, Algae, and Higher Plants. *Nat. Prod. Rep.* **1999**, *16*, 565-574.

(35) Dellas, N.; Thomas, S. T.; Manning, G.; Noel, J. P. Discovery of a Metabolic Alternative to the Classical Mevalonate Pathway. *eLife* [online] **2013**, *2*, e00672. http://elifesciences.org/content/2/e00672/article-info (accessed Dec 22, 2014).

(36) Azami, Y.; Hattori, A.; Nishimura, H.; Kawaide, H.; Yoshimura, T.; Hemmi, H. (R)-mevalonate 3-phosphate is an Intermediate of the Mevalonate Pathway in *Thermoplasma acidophilum*. *J. Biol. Chem.* [online]. DOI: 10.1074/jbc.M114.562686. Published Online: Apr 22, 2014. http://www.jbc.org/content/early/2014/04/22/jbc.M114.562686.full.pdf+html (accessed Dec 22, 2014).

(37) Poulter, C. D.; Rilling, H. C. The Prenyl Transfer Reaction. Enzymatic and Mechanistic Studies of the 1'-4 Coupling Reaction in the Terpene Biosynthetic Pathway. *Accounts Chem. Res.* **1978**, *11*, 307-313.

(38) Thulasiram, H. V.; Poulter, C. D. Farnesyl Diphosphate Synthase: The Art of Compromise between Substrate Selectivity and Stereochemistry. *J. Am. Chem. Soc.* **2006**, *128*, 15819-15823.

(39) Resh, M. D. Trafficking and Signaling by Fatty-acylated and Prenylated Proteins. *Nat. Chem. Biol.* **2006**, *2*, 584-590.

(40) Heathcock, C. H.; Finkelstein, B. L.; Aoki, T.; Poulter, C. D. Stereostructure of the Archaebacterial C_{40} Diol. *Science* **1985**, *229*, 862-864.

(41) Gunawardena, K.; Rivera, S. B.; Epstein, W. W. The Monoterpenes of *Artemisia tridentata* ssp. *vaseyana, Artemisia cana* ssp. *viscidula* and *Artemisia tridentata* ssp. *spiciformis*. *Phytochemistry* **2002**, *59*, 197-203.

(42) Giner, J-L.; Berkowitz, J. D.; Andersson, T. Nonpolar Components of the Latex of *Euphorbia peplus*. *J. Nat. Prod.* **2000**, *63*, 267-269.

(43) Zhang, A.; Amalin, D.; Shirali, S.; Serrano, M. S.; Franqui, R. A.; Oliver, J. E.; Klun, J. A.; Aldrich, J. R.; Meyerdirk, D. E.; Lapointe, S. L. Sex Pheromone of the Pink Hibiscus Mealybug, *Maconellicoccus hirsutus*, Contains an Unusual Cyclobutanoid Monoterpene. *P. Natl. Acad. Sci. USA* **2004**, *101*, 9601-9606.

(44) Bierl-Leonhardt, B. A.; Moreno, D. S.; Schwarz, M.; Fargerlund, J.; Plimmer, J. R. Isolation, Identification, and Synthesis of the Sex Pheromone of the Citrus Mealybug, *Planococcus citri (risso). Tetrahedron Lett.* **1981**, *22*, 389-392.

(45) Millar, J. G.; Midland, S. L.; Mcelfresh, J. S.; Daane, K. M. (2,3,4,4-Tetramethylcyclopentyl)methyl Acetate, a Sex Pheromone from the Obscure Mealybug: First Example of a New Structural Class of Monoterpenes. *J. Chem. Ecol.* **2005**, *31*, 2999-3005.

(46) Rilling, H. C. A New Intermediate in the Biosynthesis of Squalene. *J. Biol. Chem.* **1966**, *13*, 3233-3236.

(47) Rilling, H. C.; Epstein, W. W. Mechanism of Squalene Biosynthesis. Presqualene, a Pyrophosphorylated Precursor to Squalene. *J. Am. Chem. Soc.* **1969**, *91*, 1041-1042.

(48) Altman, L. J.; Ash, L.; Kowerski, R. C.; Epstein, W. W.; Larsen, B. R.; Rilling, H. C.; Muscio, F.; Gregonis, D. E. Prephytoene Pyrophosphate. New Intermediate in the Biosynthesis of Carotenoids. *J. Am. Chem. Soc.* **1972**, *94*, 3257-3259.

(49) Epstein, W. W.; Poulter, C. D. A Survey of Some Irregular Monoterpenes and their Biogenetic Analogies to Presqualene Alcohol. *Phytochemistry*, **1973**, *12*, 737-747.

(50) Poulter, C. D.; Marsh, L. L.; Hughes, J. M.; Argyle, J. C.; Satterwhite, D. M. ; Goodfellow, R. J.; Moesinger, S.G. Model Studies of the Biosynthesis of Non-Head-to-Tail Terpenes. Rearrangements of the Chrysanthemyl System. *J. Am. Chem. Soc.* **1977**, *99*, 3816-3823.

(51) Zheng, H.; Poulter, C. D. Stereochemical Studies of Botryococcene Biosynthesis: Analagies Between 1'-1 and 1'-3 Condensations in the Isoprenoid Pathway. *J. Am. Chem. Soc.* **1989**, *111*, 2713-2715.

(52) Poulter, C. D. Biosynthesis of Non-Head-to-Tail Terpenes. Formation of 1'-1 and 1'-3 Linages. *Acc. Chem. Res.* **1990**, *23*, 70-77.

(53) Behrens, A.; Schaeffer, P.; Bernasconi, S.; Albrecht, P. 7,11-Cyclobotyococca-5, 12, 26-triene, a Novel Botryococcene-Related Hydrocarbon Occurring in Natural Environments. *Org. Lett.* **2000**, *2*, 1271-1274.

(54) Kambara, H.; Yamada, T.; Tsujioka, M.; Matsunaga, S.; Tanaka, R.; Ali, H. I.; Wiart, C.; Yusof, M.; Hassan, H.; Hanifah, A.; Fauzi, Z. M.; Mazlan, N. H.; Jay, M.; Kunishima, M.; Akaho, E. A Study on Medicinal Plants from Malaysia Focused on *Acalypha siamensis* Oliv. *ex*Gage. Isolation and Structure of a New Tetraterpene, Acalyphaser A. *Chem. Biodivers.* **2006**, *3*, 1301-1306.

(55) Pan, J-J.; Bugni, T. S.; Poulter, C. D. Recombinant Squalene Synthase. Synthesis of Cyclopentyl Non-Head-to-Tail Triterpenes. *J. Org. Chem.* **2009**, *74*, 7562-7565.

(56) Thulasiram, H. V.; Erickson, H. K.; Poulter, C. D. Chimeras of Two Isoprenoid Synthases Catalyze All Four Coupling Reactions in Isoprenoid Biosynthesis. *Science* **2007**, *316*, 73-76.

(57) Perez-Brocal, V.; Gil, R.; Ramos, S.; Lamelas, A.; Postigo, M.; Michelena, J. M.; Silva, F. J.; Moya, A.; Latorre, A. A Small Microbial Genome: The End of a Long Symbiotic Relationship? *Science* **2006**, *314*, 312-313.

(58) Tarshis, L. C.; Yan, M.; Poulter, C. D.; Sacchettini, J. C. Crystal Structure of Recombinant Farnesyl Diphosphate Synthase at 2.6-.ANG. Resolution. *Biochemistry* **1994**, *33*, 10871-10877.

(59) Poulter, C. D. Farnesyl diphosphate synthase. A paradigm for understanding structure and function relationships in E-polyprenyl diphosphate synthases. *Phytochem. Rev.* **2006**, *5*, 17-26.

(60) Pandit, J.; Danley, D. E.; Schulte, G. K.; Mazzalupo, S.; Pauly, T. A.; Hayward, C. M.; Hamanaka, E. S.; Thompson, J. F.; Harwood, Jr., H. J. Crystal Structure of Human Squalene Synthase. *J. Biol. Chem.* **2000**, *275*, 30610-30617.

(61) Rivera, S. B.; Swedlund, B. D.; King, G. J.; Bell, R. N.; Hussey, Jr., C. E.; Shattuck-Eidens, D. M.; Wrobel, W. M.; Peiser, G. D.; Poulter, C. D. Chrysanthemyl Diphosphate Synthase: Isolation of the Gene and Characterization of the Recombinant Non-head-to-tail Monoterpene Synthase from *Chrysanthemum cinerariaefolium*. *Proc. Natl. Acad. Sci.* **2001**, *98*, 4373-4378.

(62) Christianson, D. W. Structural Biology and Chemistry of the Terpenoid Cyclases. *Chem. Rev.* **2006**, *106*, 3412-3442

(63) Structure Function Linkage Database. http://sfld.rbvi.ucsf.edu/django/superfamily/5/sequences/all/ (accessed 04/30/14). Isoprenoid Synthase Type I Superfamily.

(64) Chayet, L.; Rojas, M. C.; Cori, O.; Bunton, C. A.; McKenzie, D. C. Complexes of Bivalent Cations with Neryl and Geranyl Pyrophosphate: Their Role in Terpene Biosynthesis. *Bioorg. Chem.* **1984**, *12*, 329-338.

(65) Hosfield, D. J.; Zhang, Y.; Dougan, D. R.; Broun, A.; Tari, L. W.; Swanson, R. V.; Finn, J. Structural Basis for Bisphosphonate-mediated Inhibition of Isoprenoid Biosynthesis. *J. Biol. Chem.* **2004**, *279*, 8526-8529.

(66) Reardon, D.; Farber, G. K. The Structure and Evolution of α/β Barrel Proteins. *FASEB J.* **1995**, *9*, 497-503.

(67) Hemmerlin, A.; Rivera, S. B.; Erickson, H. K.; Poulter, C. D. Enzymes Encoded by the Farnesyl Diphosphate Synthase Gene Family in the Big Sagebrush *Artimisia tridentata* ssp. spiciformis. *J. Biol. Chem.* **2003**, *278*, 32132-32140.

(68) Chen, A.; Kroon, P. A.; Poulter, C. D. Isoprenyl Diphosphate Synthases: Protein Sequence Comparisons, A Phylogenetic Tree, and Predictions of Secondary Structure. *Protein Sci.* **1994**, *3*, 600-607.

(69) Hansson, L. O.; Mannervik, B. Use of Chimeras Generated by DNA Shuffling: Probing Structure-Function Relationships among Glutathione Transfereases. *Method Enzymol.* **2000**, *328*, 463-477.

(70) Choi, M.; Hwang, E. Y.; Kim, E.; Huh, J.; Cho, S. Identification of amino acids responsible for different GTP preferences of human glutamate dehydrogenase enzymes. *Biochem. Bioph. Res. Co.* **2008**, *368*, 742-747.

(71) Conrado, R. J.; Varner, J. D.; DeLisa, M. P. Engineering the Spatial Organization of Metabolic Enzymes: Mimicking Nature's Synergy. *Curr. Opin. Biotech.* **2008**, *19*, 492-499.

CHAPTER 2

EXPERIMENTAL RESULTS

2.1 Chimeragenesis Design

The FPPase and CPPase chimeric enzyme constructs were built using the N-terminal sequence of one enzyme up to one of the 13 crossover points, at which the sequence from the other enzyme was used to finalize the gene. The polymerase chain reaction (PCR) protocol developed for chimeragenesis directionally cloned the chimeric genes into the pET15b vector (Novagen), linking an amino-terminal hexa-histidine tag to the expressed proteins to facilitate purification by Ni^{2+} Sepharose column chromatography. Naming of the chimeras was based upon the enzyme from which the N-terminal sequence was derived, F or C for FPPase and CPPase, respectively, and crossover point at which the sequence changes to the other enzyme. For instance, the "C6" chimera consists of CPPase sequence from the N-terminus to the sixth crossover point, with the remaining sequence from FPPase.

The beginning and ends of each IS-1 fold helix marks a crossover point, with an additional point between K193 and T194G (crossover 10) to separate the helical kinked pair of helix V. The locations of the IS-1 helices were identified from an apo sagebrush FPPase crystal structure (unpublished). The set of crossover points used in this work isolate the N- and C- terminal regions, and helices and loops of the IS-1 fold, building

enzymes of varying FPPase- and CPPase-like character using homologous regions. Each chimera provides insight into how a specific structural region defines the catalytic specificities and activities of the proteins (Figure 2.1). To fully assess the influence of the enzymatic structure outside of IS-1 fold, two additional "core-scaffold" chimeras were built. The additional proteins consist of the IS-1 fold region from one enzyme (core) with the N- and C-terminal regions of the other (scaffold).

2.2 Enzymatic Product Ratio Analyses

2.2.1 Enzymatic Products Formed Under Chain Elongation Reaction Conditions

Enzymatic products formed under chain elongation reaction conditions were determined by incubating the enzymes with 500 μM DMAPP and IPP at 30 °C for 2 h (Figure 2.2, Table 2.1). The diphosphate compounds were hydrolyzed to their alcohol counterparts through the action of calf-intestinal phosphatase (CIP), and then extracted using methyl tert-butyl ether (MTBE) for structural analysis by GC-MS.

As the N- to C-terminal transformation of FPPase into CPPase progressed, preferential FPP formation (> 78%) was maintained up through chimera C5. At the C6 chimera, production of the C_{10} and C_{15} chain elongation products GPP and FPP, respectively, was evenly split. The shift to preferential GPP formation began at the C7 chimera (61% GPP), with GPPase activity peaking at the C9 chimera (~ 95% GPP). GPP sustained as the major chain elongation product through the remaining FPPase to CPPase metamorphosis (Figure 2.3).

Chimera C13 marked a significant point in the FPPase to CPPase conversion.

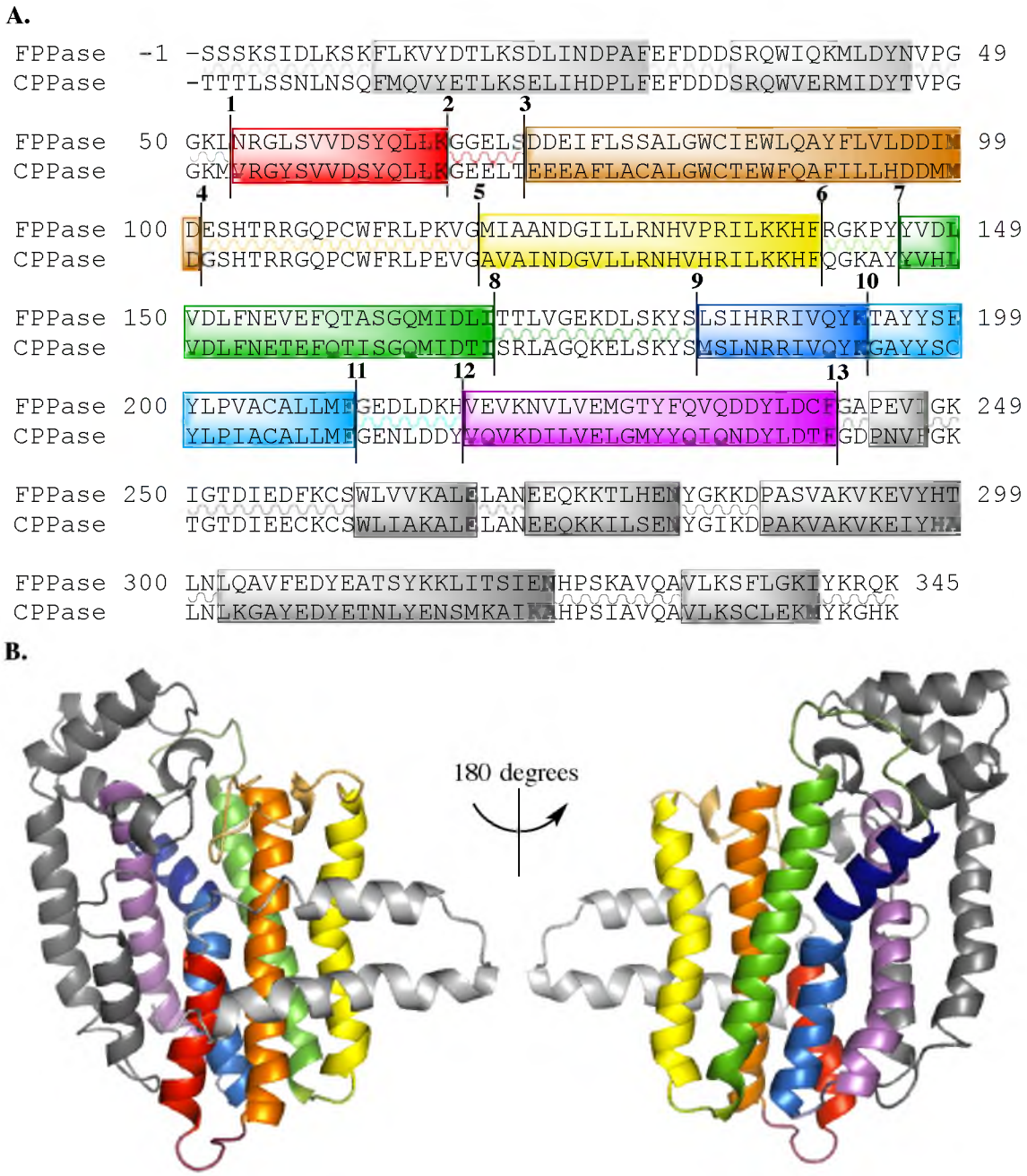
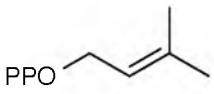
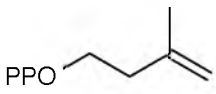


Figure 2.1. Sequential and structural representation of the parental sagebrush enzymes highlighting the crossover points for chimeragenesis **A.** Sage FPPase and CPPase amino acid sequence alignment highlighting the helices and loops of the N-terminal (light gray), IS-1 fold (colored), and C-terminal (gray) regions. The 13 crossover points used for chimeragenesis are located at the beginning and end of each IS-1 fold helix. **B.** Monomeric Sage FPPase structure showing the N-terminal, IS-1 fold, and C-terminal regions.

Substrates:

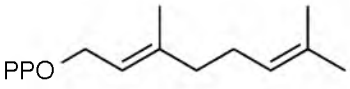


Dimethylallyl Diphosphate
(DMAPP)

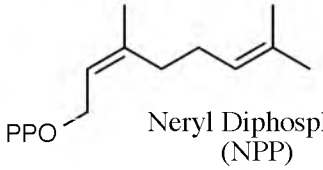


Isopentenyl Diphosphate
(IPP)

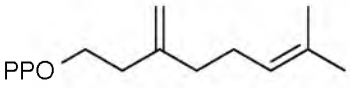
Chain Elongation / Regular Coupling Products:



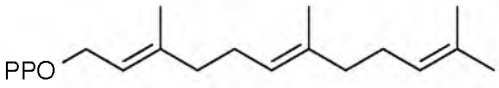
Geranyl Diphosphate
(GPP)



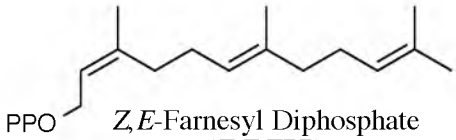
Neryl Diphosphate
(NPP)



3-Methylene GPP
(iGPP)



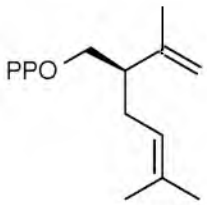
Farnesyl Diphosphate
(FPP)



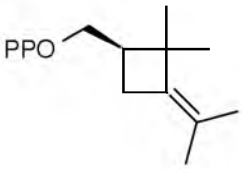
Z,*E*-Farnesyl Diphosphate
(*Z*,*E*-FPP)

Irregular Coupling Products:

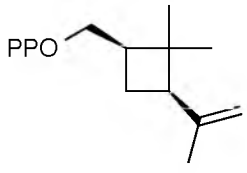
(1R,3R)-Chrysanthemyl Diphosphate
(CPP)



(R)-Lavandulyl Diphosphate
(LPP)



(R)-Maconelliyl Diphosphate
(MPP)



(1R,3R)-Planococcyl Diphosphate
(PPP)

Figure 2.2. Substrates used and products formed by FPPase, CPPase, and the chimeric enzyme constructs.

Table 2.1. Relative percentages of products formed by the FPPase, CPPase, C1-C13, F1-F13, c_F_c, and f_C_f enzymes under chain elongation incubation conditions (GPP, geranyl diphosphate; FPP, farnesyl diphosphate; CPP, chrysanthemyl diphosphate; LPP, lavandulyl diphosphate).

	GPP	FPP	CPP	LPP
FPPase	8 ± 1	90 ± 1	-	-
C1	9 ± 2	91 ± 2	-	-
C2	21 ± 1	79 ± 1	-	-
C3	6.8 ± 0.7	93.2 ± 0.7	-	-
C4	15.9 ± 0.7	84.1 ± 0.7	-	-
C5	22 ± 3	78 ± 3	-	-
C6	52 ± 6	48 ± 6	-	-
C7	61 ± 2	39 ± 2	-	-
C8	67 ± 2	32 ± 2	-	-
C9	94.5 ± 0.3	3.1 ± 0.2	-	-
C10	78 ± 2	21 ± 2	-	-
C11	72 ± 4	27 ± 4	-	-
C12	88.4 ± 0.9	10 ± 1	-	-
C13	64 ± 1	23 ± 1	2.7 ± 0.2	7.5 ± 0.2
CPPase	49 ± 1	2.0 ± 0.1	34 ± 1	11.4 ± 0.1
F1	64.9 ± 0.7	4.3 ± 0.9	17 ± 1	10.2 ± 0.6
F2	93.3 ± 0.6	4.4 ± 0.6	-	-
F3	81 ± 6	19 ± 6	-	-
F4	76 ± 2	24 ± 2	-	-
F5	67 ± 2	33 ± 2	-	-
F6	72 ± 3	28 ± 3	-	-
F7	57 ± 8	43 ± 8	-	-
F8	14 ± 2	83 ± 2	-	-
F9	8 ± 2	90 ± 2	-	-
F10	25.6 ± 0.8	71.9 ± 0.5	-	-
F11	38 ± 4	60 ± 4	-	-
F12	53 ± 3	46 ± 4	-	-
F13	21 ± 5	77 ± 5	-	-
c_F_c	11 ± 3	86 ± 3	-	-
f_C_f	70 ± 2	-	0.5 ± 0.1	25 ± 1

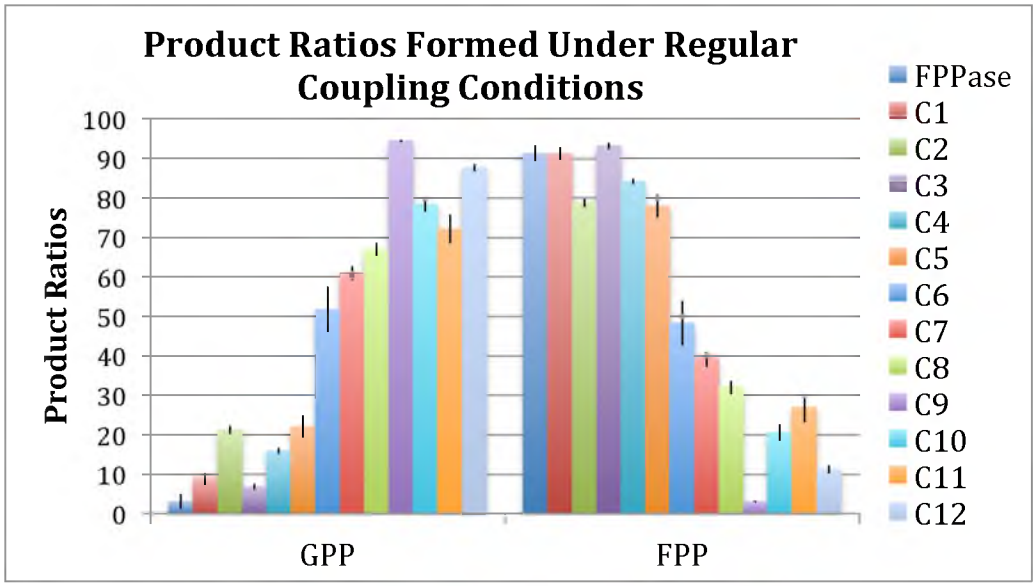


Figure 2.3. Preferential formation of FPP shifts to preferential formation of GPP along the N- to C-terminal FPPase to CPPase metamorphosis (GPP, geranyl diphosphate; FPP, farnesyl diphosphate).

Structurally, the chimera was almost entirely built from CPPase sequence, only the C-terminal region outside the IS-1 fold, was sequence from FPPase. The C13 chimera exhibited irregularly coupled products (~ 10%) under chain elongation incubation conditions, similar to CPPase (~ 45% irregular products), yet in contrast to CPPase, chimera C13 preferentially formed the branched product LPP over the cyclopropyl CPP (Figure 2.4).

The first step of the N- to C-terminal, CPPase to FPPase transformation is the F1 chimera, which is the CPPase enzyme with N-terminal sequence from FPPase. The F1 chimera also competitively catalyzed irregular and regular couplings under incubation with DMAPP and IPP. The F1 chimera preferentially catalyzes chain elongation to GPP, but similarly to CPPase, formed CPP as the major irregular product (65% GPP, 17% CPP, 10% LPP). Competitive irregular coupling ability was lost in the CPPase to FPPase

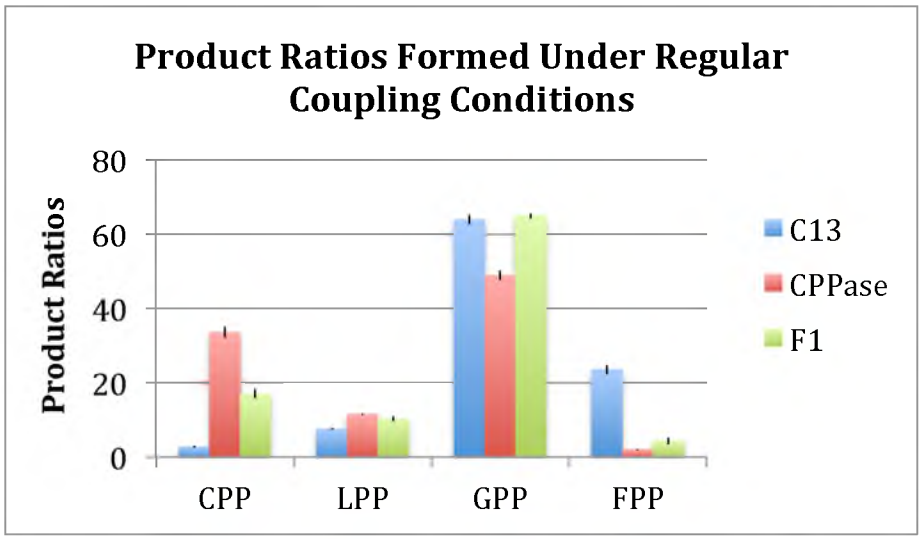


Figure 2.4. CPPase, chimera C13, and chimera F1 competitively form regularly and irregularly coupled products under regular coupling conditions (CPP, chrysanthemyl diphosphate; LPP, lavandulyl diphosphate; GPP, geranyl diphosphate; FPP, farnesyl diphosphate).

metamorphosis at the F2 chimera (93.3% GPP). In continuation along the metamorphosis, the formation of FPP was steadily reclaimed, reaching a maximum at the F9 chimera (90% FPP). Interestingly, FPP production then decreased at the F10 chimera (~ 72% FPP), and continued to decrease through the F12 chimera, where GPP and FPP production nearly equalized. The trend reversed once again, and FPP production was reclaimed at the F13 chimera (77% FPP) (Figure 2.5).

The chain elongation product ratios for the FPPase core-CPPase scaffold (c_F_c) chimera were comparable to parental FPPase (86% FPP). The CPPase core-FPPase scaffold (f_C_f) chimera resembled chimera C13, showing preferential chain elongation to GPP (70%), with LPP formed almost exclusively over CPP (25% and 0.5%, respectively).

Three additional minor products were observed in the chain elongation analyses,

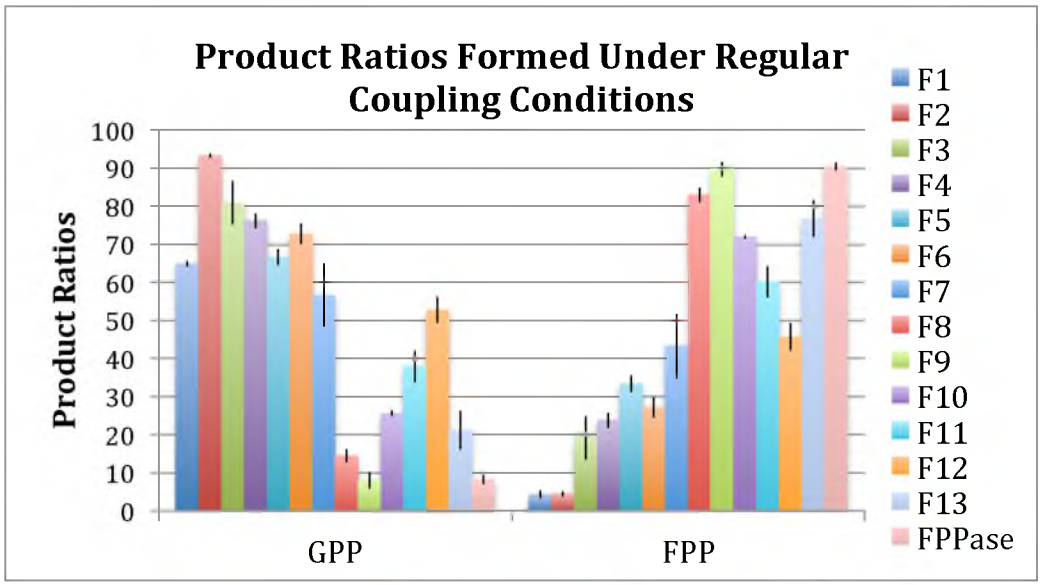


Figure 2.5. Relative formation of GPP and FPP for FPPase and the F1-F13 chimeras (GPP, geranyl diphosphate; FPP, farnesyl diphosphate).

two being the *Z*-isomers of GPP and FPP, neryl diphosphate (NPP) and *Z,E*-FPP, respectively. Generally, the *Z*-isomers were produced in low amounts (1-2.5%) by enzymes forming a large amount of their *E*-isomer counterparts. The third minor product was identified as a double-bond isomer of GPP, 3-methylene GPP (iGPP).[1] Only CPPase and the F1 chimera were observed to catalyze iGPP formation in relatively low amounts (1-2%).

2.2.2 Enzymatic Products Formed Under Irregular Terpenoid Coupling Conditions

Enzymatic formation of irregularly coupled terpenoid products was studied in a similar manner as the chain elongation products, except the enzymes were incubated with 3 mM DMAPP at 30 °C overnight (Table 2.2). Formation of irregular isoprenoids by

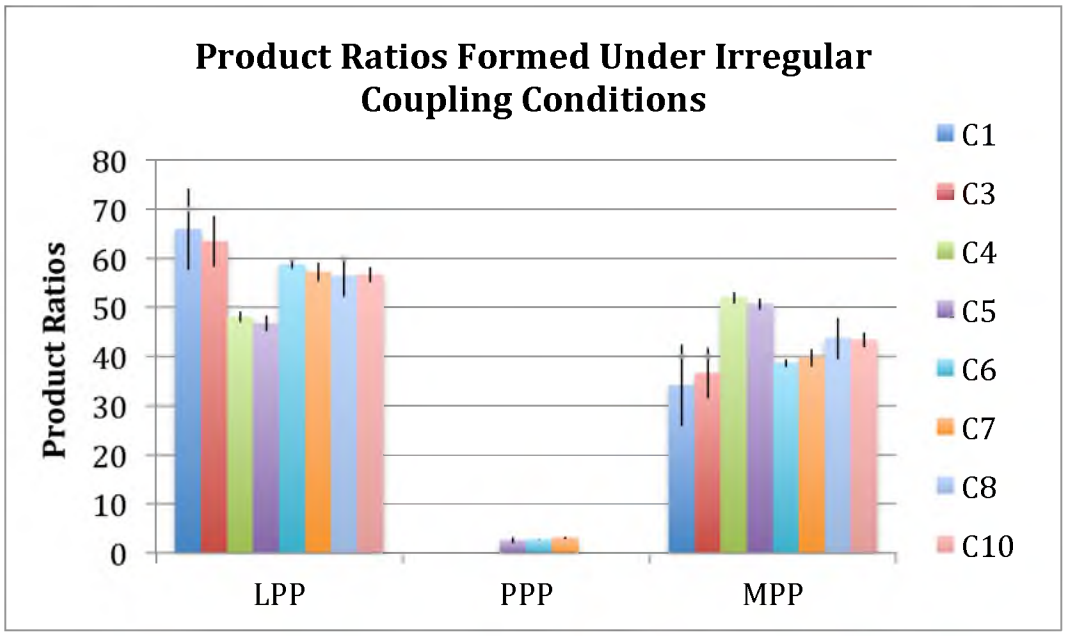
Table 2.2. Relative percentages of enzymatic products formed through the irregular coupling of two DMAPP molecules (chrysanthemyl diphosphate, CPP; lavandulyl diphosphate, LPP; maconelliyl diphosphate, MPP; planococcyl diphosphate, PPP).

	CPP	LPP	MPP	PPP
C1	-	66 ± 8	34 ± 8	-
C3	-	63 ± 5	37 ± 5	-
C4	-	48 ± 1	52 ± 1	-
C5	-	47 ± 2	50 ± 1	3 ± 0.6
C6	-	58 ± 1	39 ± 1	3 ± 0.1
C7	-	57 ± 2	40 ± 2	3 ± 0.2
C8	-	56 ± 4	44 ± 4	-
C10	-	57 ± 1	43 ± 1	-
C11	-	94 ± 1	6 ± 1	-
C12	2 ± 0.3	95 ± 1	3 ± 0.4	-
C13	27 ± 0.5	72 ± 1	1 ± 0.1	-
CPPase	77 ± 1	22 ± 0.4	1 ± 0.1	-
F1	64 ± 1	35 ± 1	1 ± 0.2	-
F8	5 ± 0.2	86 ± 1	9 ± 0.8	-
F9	4 ± 0.8	86 ± 1	10 ± 0.3	-
F10	-	84 ± 1	17 ± 1	
f_C_f	16 ± 0.7	83 ± 1	1 ± 0.2	-

FPPase, C2, F2-F7, F10-F13, and c_F_c was not detected. For the C1-C10 chimeras, LPP and MPP were the major products, generally with a preference towards formation of the branched product. The C5-C7 chimeras are unique in that they were the only enzymes that produced the cyclobutyl product planococcyl diphosphate (PPP) (Figure 2.6).

A shift in the irregular product ratios occurred at chimera C11, with selective LPP formation observed (94% LPP, 6% MPP). The first instance of CPP formation along the transformation began at the C12 chimera (1.8% CPP), and the occurrence increased at chimera C13 (27% CPP). The cyclopropanation reaction did not out-compete branching until complete transformation into CPPase (77% CPP). The F1 chimera maintained CPPase-like product ratios, and also exhibited CPP as the major irregular isoprenoid product (64% CPP). Significant CPP formation appeared at the expense of MPP, as the C12, C13, CPPase, and F1 enzymes only produced the cyclobutyl product in trace

A.



B.

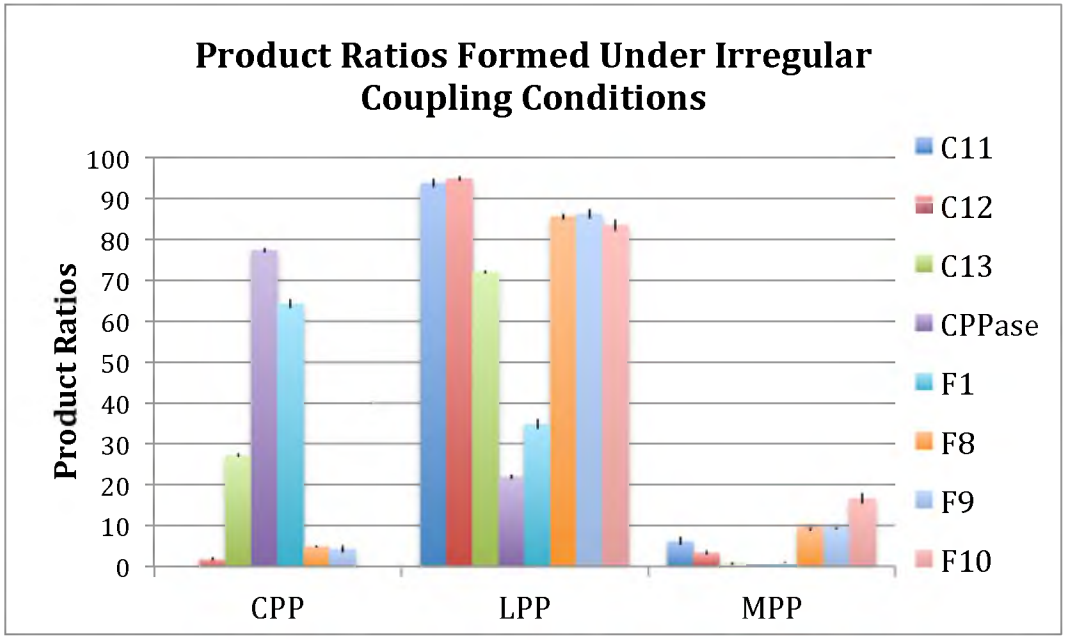


Figure 2.6. Observed product ratios for the enzymes under irregular coupling conditions. **A.** Ratios of irregularly coupled products formed by the C1, C3-C8, and C10 chimeras. **B.** Ratios of irregularly coupled products formed by the C11-C13, CPPase, F1, and F8-F9 enzymes (CPP, chrysanthemyl diphosphate; LPP, lavandulyl diphosphate; PPP, planococcyl diphosphate; MPP, maconelliyl diphosphate).

relative quantities (< 1% MPP).

In jumping from the F1 to the F8 and F9 chimeras in the CPPase to FPPase, N- to C- terminal transformation, LPP returned as the major product (~ 86% LPP), while MPP formation restored in preference to, and catalyzed in approximately 2-fold excess over, CPP (~ 4.5% CPP, ~ 9.6% MPP). As similarly observed in the chain elongation product ratios, the CPPase core-FPPase scaffold chimera showed catalytic specificity analogous to chimera C13.

2.3 Michaelis-Menten Kinetic Analyses

2.3.1 Apparent Michaelis-Menten Kinetic Values for the Chain Elongation Reaction

Rates for chain elongation were determined by thin layer chromatography (TLC) analysis of product formation in samples from single-point measurements with ^{32}P-radiolabeled IPP. Substrates and products were resolved on silica plates, and the radioactivity was visualized and quantitated by phosphorimaging (Tables 2.3 and 2.4).

Sagebrush FPPase was unrivaled in the formation of FPP, showing a catalytic efficiency relative to GPP of 2.4 (CE^{GPP}, $k_{cat}{}^{GPP}/K_M{}^{GPP,EA}$ ($s^{-1}\mu M^{-1}$)), EA indicates binding to the electron acceptor site). CE^{GPP} quickly dropped along the N- to C-terminal metamorphosis of FPPase to CPPase. Firstly, the C1 chimera showed a 5-fold decrease in CE^{GPP} relative to FPPase, with an additional 20-fold drop observed at chimera C2. The C2-like CE^{GPP} values were maintained through chimera C6.

At the C6 to C7 transition, a nearly 30-fold drop in CE^{GPP} occurred – an almost 4000-fold decrease relative to FPPase. The CE^{GPP} values over the C7-C10 chimeric

Table 2.3. Apparent Michaelis-Menten kinetic values obtained under chain elongation conditions using [^{32}P]IPP (EA indicates binding in the electron acceptor site, ED indicates binding to the electron donor site).

	$k_{cat}^{DMAPP/IPP}$ (s^{-1})	$K_M^{DMAPP,EA}$ (μM)	$K_M^{IPP,ED}$ (μM)	$k_{cat}^{GPP/IPP}$ (s^{-1})	$K_M^{GPP,EA}$ (μM)
FPPase	0.96 ± 0.09	9 ± 2	11 ± 3	2.6 ± 0.1	0.8 ± 0.3
C1	1.2 ± 0.1	4 ± 1	12 ± 5	1.5 ± 0.1	3.5 ± 0.9
C2	0.80 ± 0.07	74 ± 6	60 ± 20	0.81 ± 0.05	38 ± 7
C3	0.63 ± 0.06	200 ± 20	23 ± 7	0.22 ± 0.01	36 ± 7
C4	1.0 ± 0.1	330 ± 50	34 ± 9	0.26 ± 0.01	26 ± 4
C5	0.61 ± 0.04	110 ± 20	14 ± 3	0.23 ± 0.01	16 ± 3
C6	0.23 ± 0.02	29 ± 3	8 ± 2	0.42 ±0.03	25 ± 5
C7	0.188 ± 0.008	140 ± 20	11 ± 2	0.147 ± 0.009	240 ± 50
C8	0.132 ± 0.006	18 ± 2	14 ± 2	0.0077 ± 0.0003	680 ± 70
C10	0.64 ± 0.03	120 ± 20	59 ± 10	0.010 ± 0.001	4000 ± 800
C11	.070 ± 0.006	2000 ± 300	1300 ± 400	-	-
C12	0.110 ± 0.008	2300 ± 300	1400 ± 400	-	-
C13	0.0111 ± 0.0003	350 ± 50	1600 ± 400	-	-
CPPase	0.0051 ± 0.0002	510 ± 90	1700 ± 200	-	-
F1	0.00125 ± 0.00009	150 ± 30	18 ± 2	0.00032 ± 0.00001	6.1 ± 0.9
F8	0.0021 ± 0.0001	1000 ± 300	1000 ± 200	0.0032 (247 μM)*	-
F9	0.0091 ± 0.0004	1000 ± 200	2000 ± 300	0.0089 (247 μM)*	-
F10	0.0067 ± 0.0004	14 ± 4	11 ± 2	0.0071 ± 0.0004	7 ± 2
F11	0.0027 ± 0.0001	1100 ± 300	1600 ± 300	0.0017 (247 μM)*	-
F12	0.00177 ± 0.00006	440 ± 80	1200 ± 100	0.0031 (82 μM)*	-
F13	0.0037 ± 0.0002	400 ± 70	2800 ± 400	0.0030 (247 μM)*	-
c_F_c	0.0087 ± 0.0003	530 ± 70	3300 ± 300	0.0097 (247 μM)*	-
f_C_f	0.012 ± 0.003	290 ± 50	80 ± 20	-	-

*The k_{cat} values indicated are from the average of two experimental samples, with the [GPP] maximizing the reaction rate is indicated. $K_M^{GPP,EA}$ could not be determined under the experimental conditions for the indicated enzymes due to the relatively low GPP and high IPP concentrations required to maximize the reaction rate (turnover at low [GPP] and high [IPP] was below the detection limit of phosphorimaging techniques).

Table 2.4. Catalytic efficiencies of the enzymes observed for the chain elongation reaction (EA indicates binding to the electron acceptor site, ED indicates binding to the electron donor site).

	$k_{cat}^{DMAPP/IPP}/K_M^{DMAPP,EA}$ ($s^{-1}\,\mu M^{-1}$)	$k_{cat}^{DMAPP/IPP}/K_m^{IPP,ED}$ ($s^{-1}\,\mu M^{-1}$)	$k_{cat}^{GPP}/K_M^{GPP,EA}$ ($s^{-1}\,\mu M^{-1}$)
FPPase	0.1	0.087	3.3
C1	0.3	0.1	0.43
C2	1.1×10^{-2}	1.3×10^{-2}	2.1×10^{-2}
C3	3.2×10^{-3}	2.7×10^{-2}	6.1×10^{-3}
C4	3.0×10^{-3}	2.9×10^{-2}	1.0×10^{-2}
C5	5.5×10^{-3}	4.4×10^{-2}	1.4×10^{-2}
C6	7.9×10^{-3}	2.9×10^{-2}	1.7×10^{-2}
C7	1.34×10^{-3}	1.7×10^{-2}	6.13×10^{-4}
C8	7.3×10^{-3}	9.4×10^{-3}	1.1×10^{-5}
C10	5.3×10^{-3}	1.1×10^{-2}	2.5×10^{-6}
C11	3.5×10^{-5}	5.4×10^{-5}	-
C12	4.8×10^{-5}	7.9×10^{-5}	-
C13	3.2×10^{-5}	6.9×10^{-6}	-
CPPase	1.0×10^{-5}	3.0×10^{-6}	-
F1	8.3×10^{-6}	6.9×10^{-5}	5.2×10^{-5}
F8	2.1×10^{-6}	2.1×10^{-6}	-
F9	9.1×10^{-6}	4.6×10^{-6}	-
F10	4.8×10^{-4}	6.1×10^{-4}	1×10^{-3}
F11	2.5×10^{-6}	1.7×10^{-6}	-
F12	4.0×10^{-6}	1.5×10^{-6}	-
F13	9.6×10^{-6}	1.3×10^{-6}	-
c_F_c	1.6×10^{-5}	2.6×10^{-6}	-
f_C_f	4.1×10^{-5}	1.5×10^{-4}	-

enzyme transition continually decreased to the point where chimera C10 produced GPP at more than 2000-fold greater catalytic efficiency than its C_{15} counterpart. The C11-C13 chimeras, as well as CPPase, did not produce FPP under the kinetic experimental conditions.

The ability to chain elongate to farnesyl diphosphate was immediately recovered along the N- to C-terminal, CPPase to FPPase metamorphosis, requiring only the

replacement of the N-terminal region of CPPase with FPPase. The F1 chimera appeared to bind GPP and IPP relatively well ($K_M^{GPP,EA}$ of 6.1 μM, $K_M^{IPP,ED}$ of 18 μM; ED indicates binding to the electron donor site), although substrate turnover was very low ($k_{cat}^{GPP/IPP}$ of 3.2×10^{-4} s^{-1}) and remained as such throughout the entire CPPase to FPPase transformation.

The kinetics regarding chain elongation to GPP mirrored the results observed with its C_{15} counterpart. FPPase and the C1 chimera were the most catalytically efficient GPP producers. At chimera C2 an approximately 10-fold decrease in catalytic efficiency relative to DMAPP and IPP occurred (CE^{DMAPP}, $k_{cat}^{DMAPP/IPP}/K_M^{DMAPP,EA}$ ($s^{-1}\mu M^{-1}$); CE^{IPP}, $k_{cat}^{DMAPP/IPP}/K_M^{IPP,ED}$ ($s^{-1}\mu M^{-1}$), respectively), with another 10-fold drop in CE^{DMAPP} observed at the C3 chimera. C3-like CE^{DMAPP} and CE^{IPP} levels were maintained along the FPPase to CPPase chimeric enzyme transformation up through chimera C10. A drastic change occurred at the C10 to C11 chimeric enzyme transition, where CE^{DMAPP} and CE^{IPP} decreased over 150-fold. The relatively low-level, C11-like production of GPP was essentially maintained throughout the remainder of the CPPase to FPPase transformation.

2.3.2 Apparent Michaelis-Menten Kinetic Values for the Irregular Coupling Reaction

All enzymes turning over two DMAPP molecules exhibited low catalytic rates and high K_M^{DMAPP} relative to CPPase and C13 (Table 2.5). CPPase and C13 were comparable in their irregular coupling ability, averaging a catalytic efficiency of 5.0×10^{-6} $s^{-1}\mu M^{-1}$. The K_M^{DMAPP} values obtained under chain elongation and irregular coupling conditions significantly differed for chimera C13 and CPPase. As a result, it is concluded

Table 2.5. Rates of irregular product formations observed at 6 mM DMAPP spiked with [^{32}P]DMAPP for the C1, C3-C8, C10-C12, F1, F8-F10, and f_C_f chimeras, and the apparent Michaelis-Menten parameters of CPPase and the C13 chimera (ED indicates binding in the electron donor site).

	k_{cat} (6 mM DMAPP, s^{-1})
C1	3.71×10^{-4}
C3	6.92×10^{-4}
C4	5.02×10^{-4}
C5	1.07×10^{-3}
C6	6.58×10^{-4}
C7	9.50×10^{-4}
C8	3.95×10^{-4}
C10	9.60×10^{-4}
C11	8.55×10^{-4}
C12	5.94×10^{-4}
F1	1.89×10^{-3}
F8	3.17×10^{-5}
F9	2.97×10^{-4}
F10	6.34×10^{-4}
f_C_f	1.78×10^{-3}

	k_{cat}^{DMAPP} (s^{-1})	$K_M^{DMAPP, ED}$ (μM)	$k_{cat}^{DMAPP}/K_M^{DMAPP,ED}$ ($s^{-1}\mu M^{-1}$)
C13	0.0109 ± 0.0005	2100 ± 200	5.2×10^{-6}
CPPase	0.0195 ± 0.0008	4100 ± 400	4.8×10^{-6}

respectively, under regular coupling conditions are representative of DMAPP binding to the electron acceptor site of the enzymes, consistent with the chain elongation reaction,whereas the K_M^{DMAPP} values of 2100 μM and 4100 μM for C13 and CPPase obtained under irregular coupling conditions exemplify DMAPP binding to the electron donor site.

2.4 References

(1) Pan, J-J.; Poulter, C. D. University of Utah, Salt Lake City, UT. Unpublished work, 2012.

CHAPTER 3

DISCUSSION

3.1 Introduction

Amino acid residues in the five conserved regions of *E*-chain elongation enzymes form the active sites of sagebrush FPPase and CPPase with few exceptions (Table 3.1). Four milestones along the N- to C-terminal transformation of FPPase to CPPase occur. First, increases of $K_M^{DMAPP,EA}$, $K_M^{IPP,ED}$, and $K_M^{GPP,EA}$ were observed at the C1 to C2 chimeric enzyme transition. The second milestone is marked by the transitioning of chain elongation activity from preferential FPP (C_{15}) formation to GPP (C_{10}) at chimera C10 due to an increase of $K_M^{GPP,EA}$. The third milestone appears at the C10 to C11 transition, where a decrease in $k_{cat}^{DMAPP/IPP}$, an increase in $K_M^{DMAPP,\ EA}$, and an increase in $K_M^{IPP,ED}$ occur. Finally, the fourth milestone involves the appearance of significant irregular coupling ability at chimera C13, indicating the appearance of a CPPase-like allowance of DMAPP binding to the electron donor site.

Along the N- to C-terminal conversion of CPPase to FPPase, selective binding of IPP to the electron donor site, and thus catalytic propensity towards the chain elongation reaction, was regained at the F1 chimera, followed by the loss of competitive irregular coupling ability at chimera F2. Interestingly, an FPPase-like catalytic efficiency was not restored at any point in the CPPase and FPPase transition.

Table 3.1. Direct and indirect residue-substrate interactions observed in molecular models of sagebrush FPPase and CPPase (IPP bound in the electron donor site, DMASPP bound in electron acceptor site). Unique interactions resulting from a mutation are bolded.

FPPase Residue-Substrate Contacts	CPPase Residue-Substrate Contacts
K51-IPP	K51-IPP
N53-water-IPP	
R54-IPP	R54-IPP
R54-water-IPP	R54-water-IPP
E86-water-IPP	E86-water-IPP
Q89-IPP	Q89-IPP
D96-Mg^{2+}-DMASPP	D96-Mg^{2+}-DMASPP
D100-Mg^{2+}-DMASPP	D100-Mg^{2+}-DMASPP
R105-DMASPP	R105-DMASPP
R105-water-IPP	R105-water-IPP
R106-IPP	R106-IPP
R106-water-IPP	R106-water-IPP
	Q164-water-DMASPP
D167-water-DMASPP	
	R171-water-DMASPP
K193-water-DMASPP	K193-water-DMASPP
	Y231-IPP
	N235-DMASPP
D235-Mg^{2+}-DMASPP	N235-Mg^{2+}-DMASPP
D239-Mg^{2+}-DMASPP	D239-Mg^{2+}-DMASPP
D239-water-DMASPP	
K249-DMASPP	K249-DMASPP
D253-Mg^{2+}-DMASPP	D253-Mg^{2+}-DMASPP
D253-water-DMASPP	D253-water-DMASPP
	E256-water-DMASPP
R343-water-IPP	
K345-water-IPP	K345-water-IPP

Substrate-bound molecular models of the parental enzymes and select chimeric constructs were built using an apo C3 structure (Brookhaven pdb 4KK2) and an *E. coli* FPPase structure binding IPP and an unreactive substrate analogue of DMAPP (Brookhaven pdb 1RQI, thioDMAPP or DMASPP). The *E. coli* structure was used as a template to morph the C3 structure to the catalytically relevant closed conformation, and to model IPP and DMASPP bound in the active site. The molecular models provide structural insights into how particular mutations alter the active site topologies of the enzymes, and therefore their catalytic specificities and efficiencies. The mutations correlating with events in the chimeric enzyme transitions, and their structural implications, are discussed in this chapter.

3.2 Structural Implications of the N-terminal Region in the Binding and Turnover of Isoprenoid Substrates

As exemplified by the C1 and F1 chimeras, the N-terminal portion outside the IS-1 fold of the sagebrush enzymes influences their activities and specificities. FPPase N-terminal sequence appeared to promote GPP and selective IPP binding capabilities of the enzymes, and thus their propensities toward the chain elongation reaction. Structurally, the N-terminal region of the sagebrush enzymes consist of two α-helices running perpendicular to the IS-1 fold, proximal to helices I, II, and III from the same monomer (A), and to helix V of the other monomer (B).

In FPPase and CPPase molecular models, residues from monomer A in the loop between the two N-terminal helices form hydrogen bonding interactions with residues in helix Va(B) near the fourth conserved region of E-chain elongation enzymes (E30-

R188(B), E30-Q191(B), D32-R188(B), and D33-Y192(B)). Residue K193 from the fourth conserved region assists in binding the allylic substrate to the electron acceptor site of the enzymes. Mutations of the N-terminal region may produce structural differences affecting the cross-monomer hydrogen bonding interactions, which in turn may have a cascade effect altering the position and dynamics of K193 in the electron acceptor site. This alteration of the electron acceptor site may be responsible for the $K_M^{GPP, EA}$ increase observed with the FPPase to chimera C1 transition.

The L52M FPPase to CPPase mutation is located in the N-terminal and first conserved *E*-chain elongase region of the enzymes, which form a portion of the electron donor site. Although molecular models show the L52M residues pointing away from substrate, their size difference may create unique steric interactions influencing the position of the nearby K51, N53V, and R54 residues, which do point toward and interact with substrate bound to the electron donor site. The positioning of residues of the first conserved region is a likely determinant in the selective binding of IPP to the electron donor site. In this, it is possible the steric influence of the L52M mutation is influential in the $K_M^{DMAPP,ED}$ increase and $K_M^{IPP,ED}$ decrease associated with the CPPase to F1 chimera transition.

3.3 The Loss of Catalytic Activity at the C2 Chimera

The transformation from the C1 to the C2 chimera was marked by a significant loss in catalytic efficiency for formation of chain elongation products, greatly in part from $K_M^{DMAPP,EA}$, $K_M^{GPP,EA}$, and $K_M^{IPP,ED}$ increases. An N53V FPPase to CPPase active site mutation is associated with chimera C2. As mentioned earlier, N53 indirectly coordinates

IPP in the electron donor site through a water molecule, accompanying the aforementioned K51 and R54 as the residues from the first conserved region responsible for binding substrate to the electron donor site (Figure 3.1). The N53V mutation may decrease the affinity of IPP for the site simply due to the loss of a residue-substrate polar interaction, in addition to potential steric and dynamic affects.

3.4 The Transformation from FPPase to GPPase Activity at the C10 Chimera

Enzymes chain elongating to FPP must accommodate the C_{10} hydrophobic chain of GPP. The pocket binding the aliphatic chain of GPP is formed by helices II, IV, and the Va,b kinked pair of the sagebrush IS-1 fold, and to a lesser degree helix III(B). The GPP binding pocket is altered along the N- to C-terminal, FPPase to CPPase transformation in a manner that increased $K_M^{GPP,EA}$, to the point that GPP binding was all but lost at the C10 chimera. The increase of $K_M^{GPP,EA}$ along the chimeric enzyme transition appears to be the result of a synergistic effort from several mutations involving the helices forming the GPP binding pocket.

The decline from C2-like CE^{GPP} began at the C7 chimera, which exhibited an approximately 10-fold increase in $K_M^{GPP,EA}$ and 3-fold loss in $k_{cat}^{GPP/IPP}$ relative to chimera C6. The C6-C7 chimeric transition is associated with an R141Q FPPase to CPPase mutation. R141 (in the helix III-IV loop region) appears to form two cross-monomer polar contacts, hydrogen bonding with D148(B) and D152(B) of helix IV(B) – an interaction lost with the R141Q mutation. At the C8 chimera the rate of FPP formation further decreased and $K_M^{GPP,EA}$ further increased. Two mutations correlate with the C7-

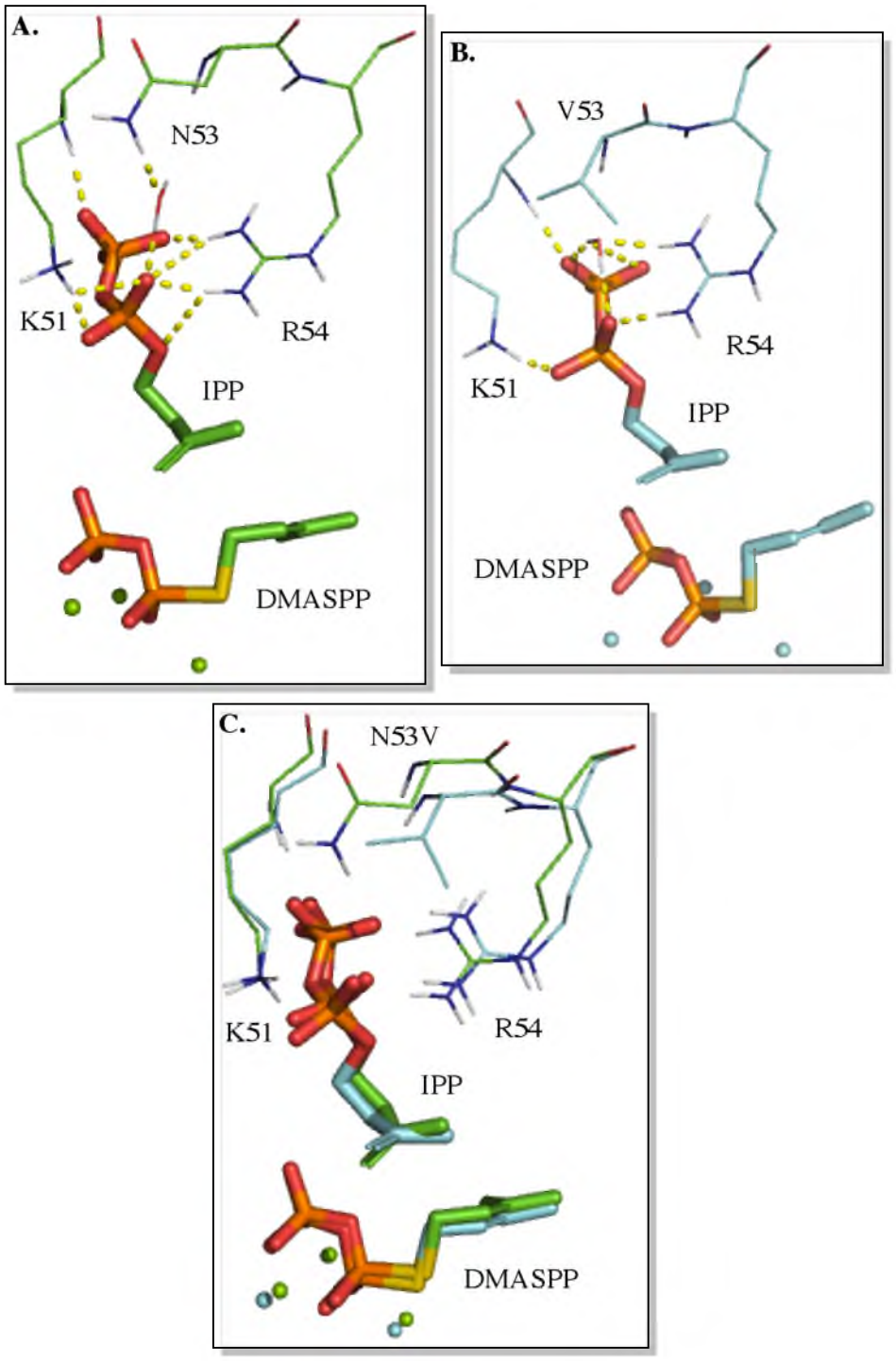


Figure 3.1. Images highlighting residues K51, N53V, and R54 of the first conserved region of E-chain elongation enzymes, which assist in binding substrate to the electron donor site. Magnesium ions are shown. The hydrogen-bonding interactions (yellow dashes) observed in FPPase (**A**, green) and CPPase (**B**, cyan) between the residues, an active site water, and IPP. **C.** Structural overlap of FPPase and CPPase.

C8 transition, A161I and L168T, which appear to alter steric interactions between helix IV and helices III and II(B), respectively. The A161I and L168T mutations surround the third conserved E-chain elongation region, which contain residues (Q164 in CPPase and C10, D167 in FPPase) that participate in binding allylic substrate to the electron acceptor site.

Along the C8 to C10 transition, GPPase activity solidified, with C10 exhibiting a more than 2000-fold catalytic efficiency toward GPP production over FPP, and doing so with a $k_{cat}^{DMAPP/IPP}$ comparable to FPPase. Two active site changes are associated with the C8 to C10 chimeric transition, T171R and H186N. Molecular models reveal that R171 of the C10 chimera inserts into the active site and participates in binding allylic substrate. In FPPase, H186 forms a polar interaction with D236 of the second DDxxD motif, perhaps providing structural stability to the electron acceptor site in a way conducive to binding GPP. FPPase and the C10 chimera show unique active site topologies revolving around D167 and the T171R mutation. In FPPase, D167 forms structural hydrogen bond interactions with K258 and K193, whereas in C10 residue D167 interacts with R171 and K193 (Figure 3.2).

It is known that mutations of the GPP binding pocket can introduce or remove steric interferences defining the length of chain an enzyme can bind, and therefore determine the chain length of the final product.[1,2] Comparison of FPPase and C10 molecular models reveal no obvious reason for the loss of GPP binding. Instead, it appears that a culmination of structural components define the nature of the GPP binding pocket, and thus the catalytic selectivity of a sagebrush enzyme.

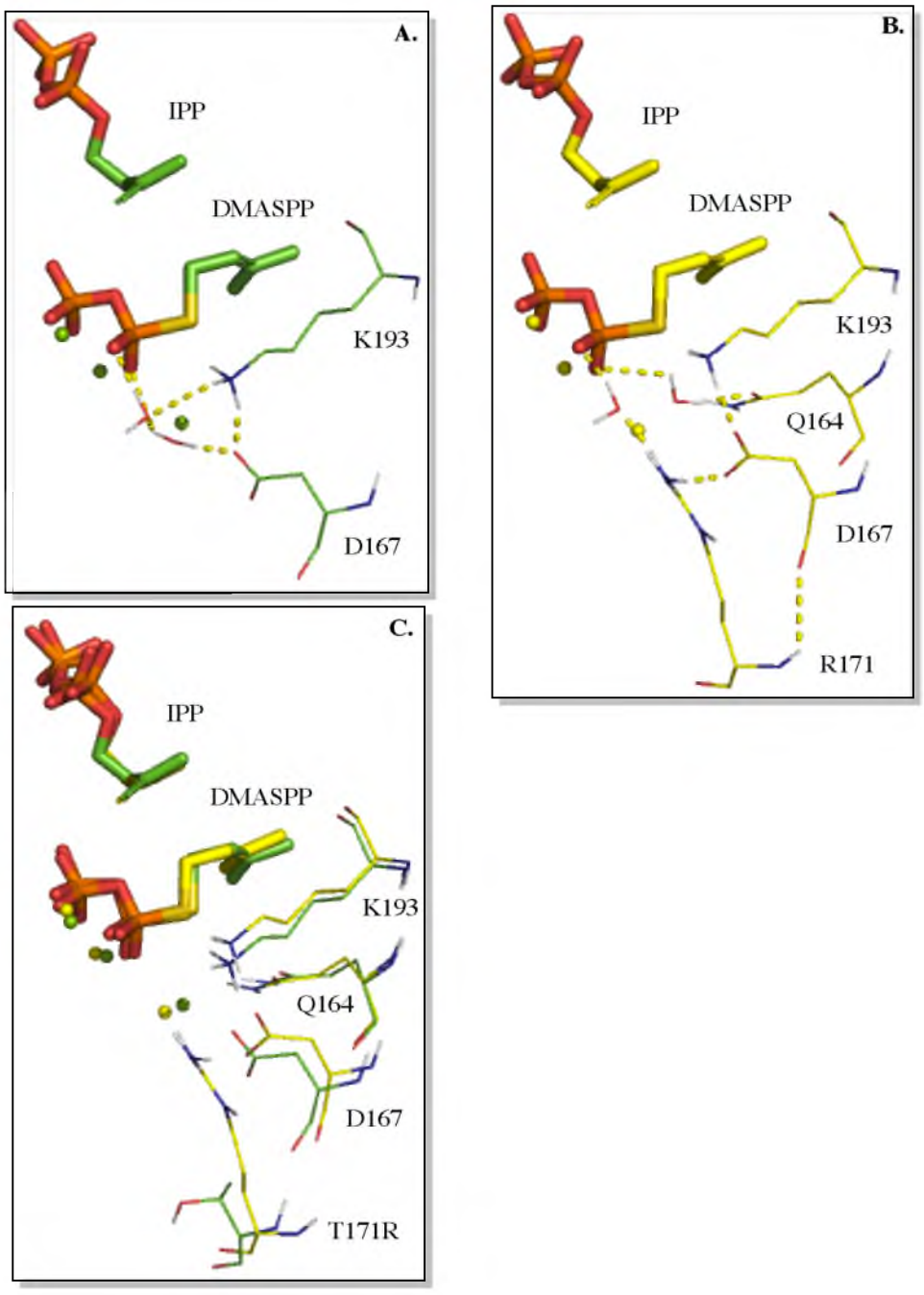


Figure 3.2. Images illustrating the structural differences between the electron acceptor binding site of FPPase (**A**, green) and the C10 chimera (**B**, yellow) due to the T171R mutation (hydrogen bonds indicated by dashed yellow lines). **C.** Structural overlap of FPPase and chimera C10. Mg^{2+} ions and relevant active site waters are shown.

3.5 The Significance of the T194G Mutation

The "KT motif" is composed of residues K193 and T194G (FPPase to CPPase mutation), which reside in the fourth conserved region among *E*-prenyl chain elongation enzymes. The fourth conserved region marks the "kink" in the Va,b helical kinked pair in the sagebrush IS-1 fold. FPPase and CPPase molecular models show K193 indirectly binding DMASPP to the electron acceptor site through water. Exploration of the role of the KT motif prompted the placement of the tenth chimeragenesis crossover point at its heart, effectively isolating the T194G mutation in the C10-C11 chimera transition.

In the molecular model of FPPase T194 forms a hydrogen bond with G228. The interaction links the Va,b helical kinked pair to helix VI, and therefore essentially the fourth and fifth conserved regions. The fifth conserved region contains the second DDxxD motif, partially responsible for binding the electron accepting allylic substrate through Mg^{2+} mediated interactions. The T194-G228 contact appears to play a significant role in the ability of the sagebrush enzymes to bind and condense terpenes, as exhibited by an approximately 10-fold reduction in $k_{cat}^{DMAPP/IPP}$ and 20-fold increase in $K_M^{DMAPP,EA}$ and $K_M^{IPP,ED}$ at the C10 to C11 transition. It appears the T194-G228 interaction provides a structural anchor to the active site, assisting to position the fourth and fifth conserved regions in a manner essential to maintaining the catalytic efficiency of the enzymes (Figure 3.3).

3.6 The Influence of the F231Y and D235N Mutations

The F231Y and D235N amino acids reside in the fifth conserved region of *E*-chain elongases. These active site mutations correlate to the C13 chimera in the N- to C-

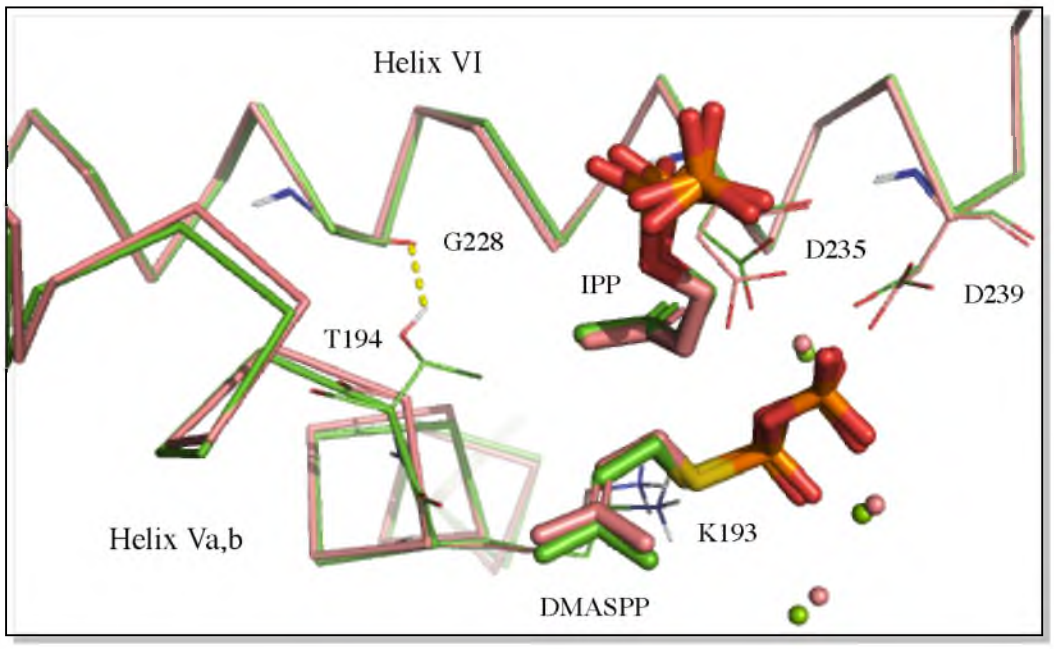


Figure 3.3. FPPase (green) and chimera C11 (salmon) structural overlaps highlighting the T194-G228 hydrogen bond (yellow dashes) in FPPase that is lost at C11 due to a T194G mutation. Also shown is K193 of the KT motif, residues D235 and D239 of the second DDxxD motif, and magnesium ions.

terminal conversion of FPPase to CPPase – the chimeragenesis point at which DMAPP competes with IPP in binding the electron donor site.

As revealed by molecular models, the F231Y mutation sterically and polarly interacts with the diphosphate moiety of substrate bound to the electron donor site. The D235N mutation appears to introduce a steric barrier in the electron donor site near the methylene carbon of IPP from the *re* face. This steric interaction forces IPP to rotate along the C2-C3 bond axis and adopt a smaller C1-C2-C3-C4 (methylene carbon as C4) dihedral angle (FPPase, 73.5°; CPPase, 53°). Interestingly, the F231Y and D235N mutations coerce the aliphatic portion of IPP into a more planar, DMAPP-like conformation – a sign very telling of their significance in creating an electron donor site suitable for DMAPP binding (Figure 3.4).

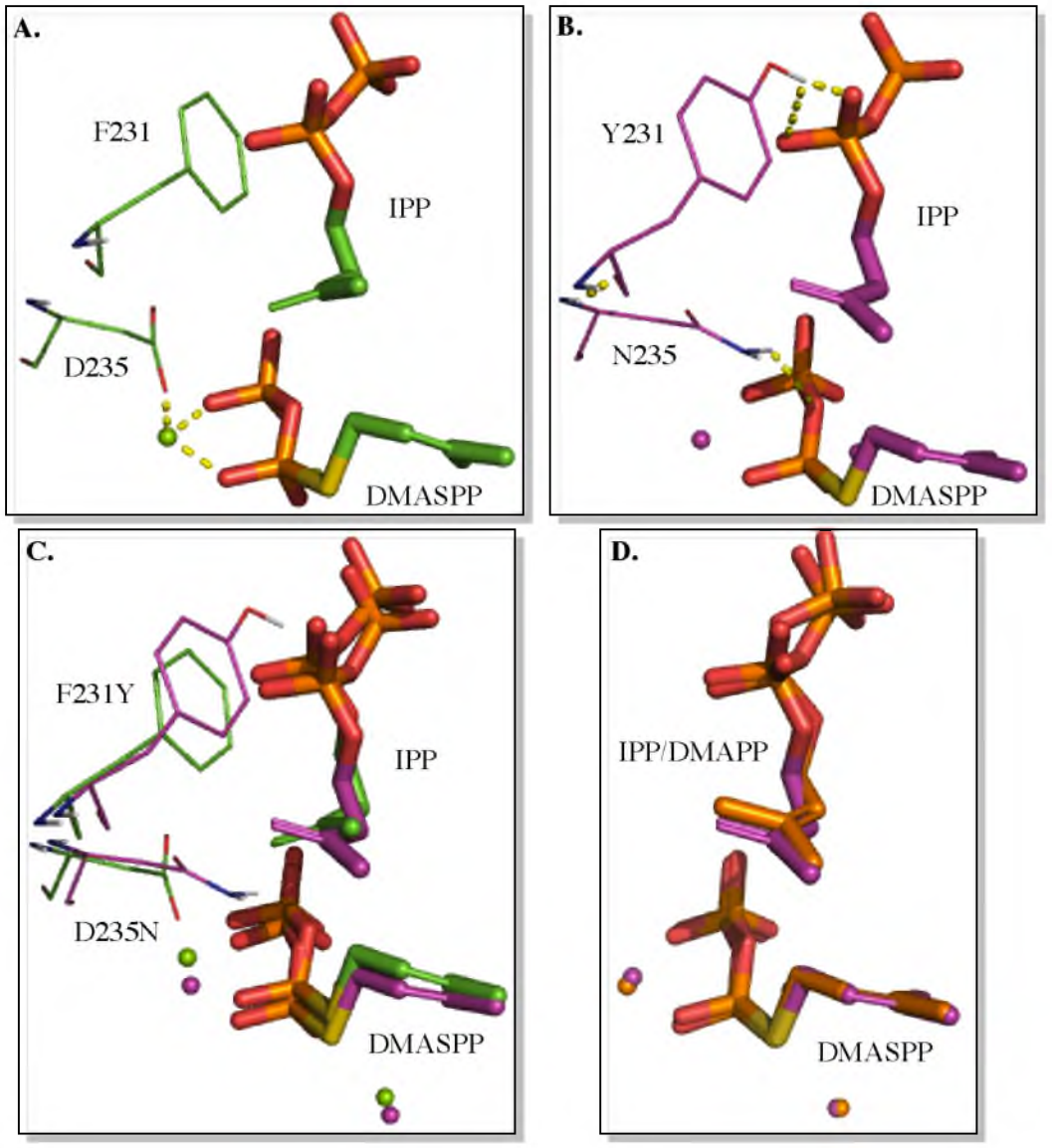


Figure 3.4. Images highlighting the influence of the F231Y and D235N mutations on the active site of the enzymes. **A.** Residues F231 and D235 in FPPase (green). **B.** Residues Y231 and N235 in the C13 chimera (magenta). **C.** Structural overlap of FPPase and chimera C13 illustrating the steric effects the mutations have on IPP positioning and conformation. **D.** C13 structural overlaps with IPP (magenta) and DMAPP (orange) bound in the electron donor site. Magnesium ions are shown.

3.7 The CPPase and C13 Enzymes – Preferential CPP versus LPP Formation

A common mechanism is utilized in the enzymatic catalysis of the fundamental irregular couplings.[2] Initiation of the irregular coupling reaction begins with the diphosphate moiety from DMAPP bound in the electron acceptor site ($DMAPP^{EA}$) leaving to form the dimethylallyl cationic intermediate (DMA^+). The DMA^+ intermediate then electrophilically alkylates the double bond of DMAPP bound in the electron donor site ($DMAPP^{ED}$), forming a protonated cyclopropyl chrysanthemyl cationic intermediate (C^+). C^+ lies at the first branch point in the irregular coupling mechanism, where the flux between CPP and LPP formation is regulated. The C^+ intermediate can either be quenched by proton extraction (presumably by the well-positioned leaving diphosphate) to form CPP, or undergo cationic rearrangement to form the branched, lavandulyl cation (L^+). The L^+ intermediate may further rearrange to form cationic cyclobutyl intermediates, or be quenched by proton extraction to form LPP. Notably, molecular models of the C13 chimera indicate that the nonleaving diphosphate (from $DMAPP^{ED}$) is better positioned to quench L^+ than the leaving diphosphate (2.6 Å and 5.1 Å, respectively) (Figure 3.5).

The ratio of CPP and LPP formation in the sagebrush enzymes depends upon the rate of proton extraction from C^+ to form CPP versus the rate of C^+ rearrangement to L^+, respectively (Scheme 3.1).[3] In the case of CPPase, proton extraction from C^+ occurs more quickly than rearrangement to L^+; conversely, rearrangement to L^+ occurs more quickly in the C13 chimera. The relative rates of proton extraction from, and rearrangement of C^+ partially depend upon the ability of the enzymes to stabilize the C^+ and L^+ cationic

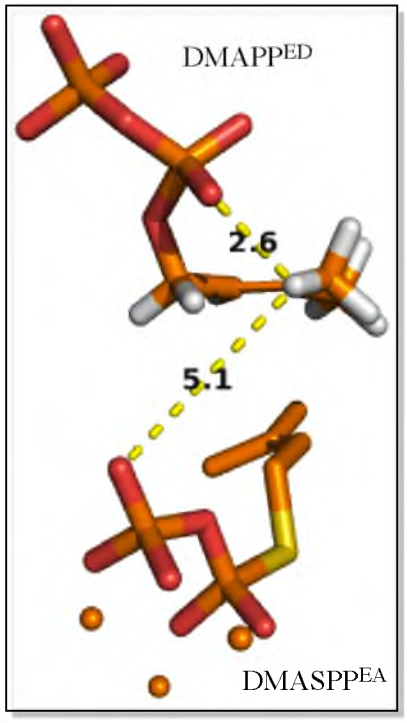
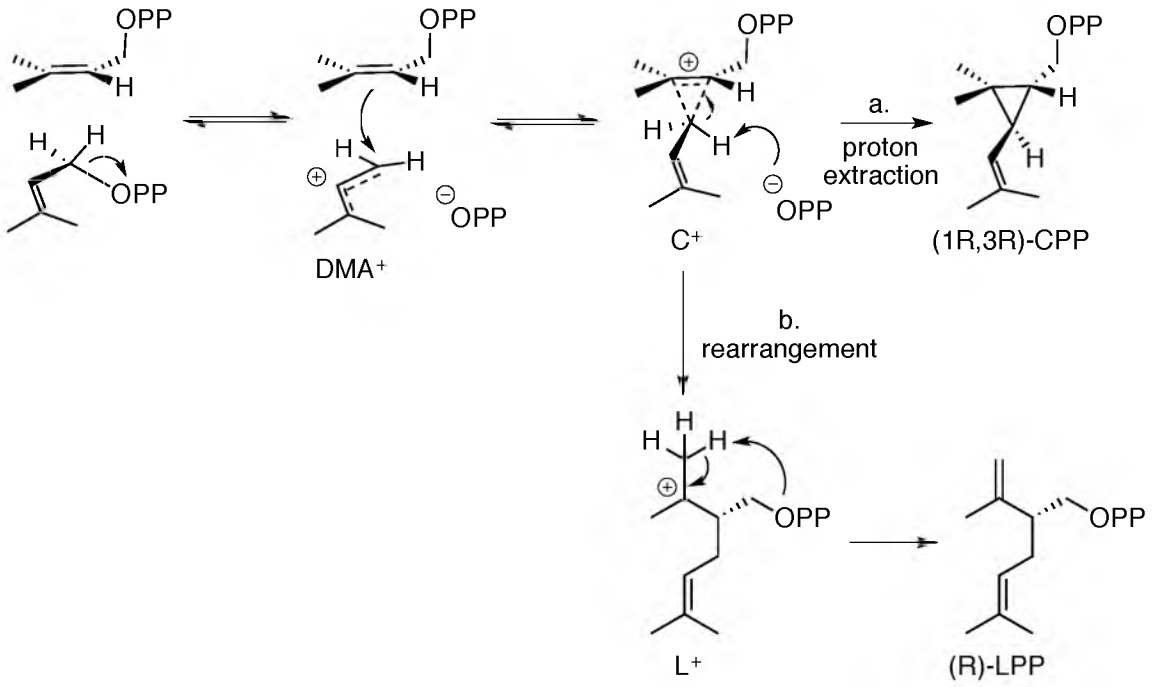


Figure 3.5. The nonleaving diphosphate from DMAPP^{ED} is better positioned at 2.6 Å to quench the lavandulyl cationic intermediate than the leaving diphosphate from DMAPP^{EA} at 5.1 Å in chimera C13.

intermediates through enzyme-substrate polar interactions.[4] CPPase and C13 differ by three active site residues – F247I, E256D, and G343R (mutations written CPPase to C13). The F247I mutation does not appear to have a significant role. In contrast, R343 in the C13 chimera is positioned to polarly interact with the diphosphate of DMAPP^{ED}, which, in turn, is positioned to provide proximal negative charge to the stabilization of L^+. In this, R343 may limit the movement of the nonleaving diphosphate, keeping it better positioned to stabilize the branched carbocation, and thus promote its formation.

If the leaving diphosphate is responsible for quenching C^+, and therefore CPP formation, positioning of the diphosphate relative to C3-C^+ is important in defining the rate of proton extraction. An inversion of symmetry at C1-DMAPP^{EA} occurs upon the



Scheme 3.1. Catalytic sequence employed by the sagebrush IS-1 enzymes in the formation of chrysanthemyl diphosphate (CPP) and lavandulyl diphosphate (LPP). The ratio of CPP versus LPP formation is dependent upon the rates of proton extraction from the chysanthemyl cation (C^+) versus rearrangement of C^+ to the lavandulyl cation (L^+).

formation of C^+, pushing the hydrogens to the same face as the leaving diphosphate. The closer the leaving diphosphate remains to the carbon it was once bonded with, the more efficiently it can extract the newly proximal proton from C3-C^+ and drive the mechanistic flux towards CPP formation. In accordance with the hypothesis, molecular models reveal that the leaving diphosphate is closer to the C^+ quenching proton in CPPase than C13 (2.7 Å versus 3.5 Å, respectively) (Figure 3.6).

The CPPase and C13 molecular models also reveal the significance of the E256D mutation in altering the electron acceptor site. In CPPase, E256 indirectly coordinates the leaving diphosphate through two water molecules, whereas D256 in the C13 chimera forms residue-residue polar contacts with T252, K258, and K285. Apparently the active

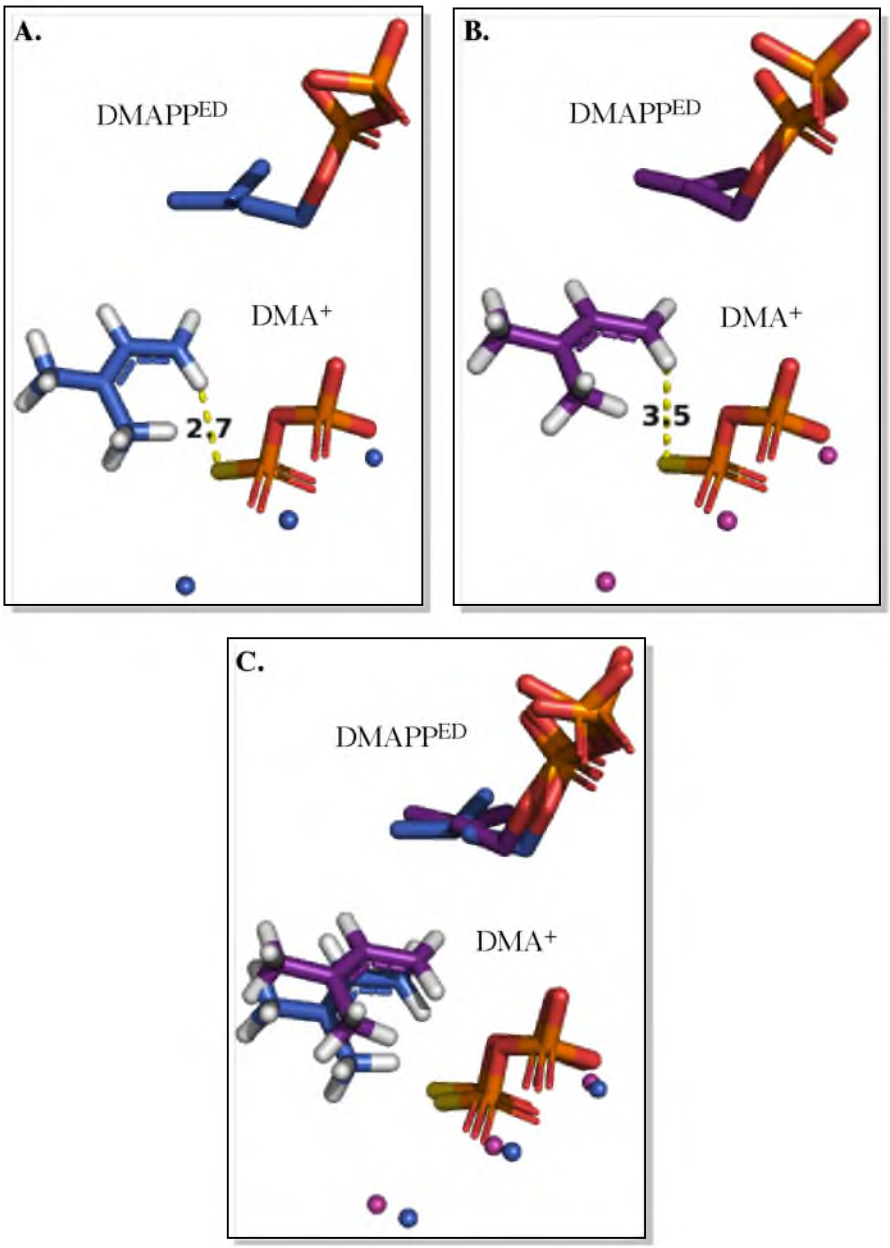


Figure 3.6. Images highlighting the distance between leaving diphosphate and the proton extracted to quench the cyclopropyl cationic intermediate (C^+) to form chrysanthemyl diphosphate (CPP) in the CPPase (blue, 2.7 Å) and C13 (magenta, 3.5 Å) enzymes. Magnesium ions are shown.

site polar interactions involving K258 depend upon whether glutamate or aspartate occupies position 256. Instead of hydrogen bonding with D256, as observed in chimera C13, K258 in CPPase forms polar residue-residue contacts with D167, T168, and the backbone of R171, effectively linking the C-terminal and third conserved E-chain elongase regions. Residue Q164 of the third conserved region sterically interacts with the aliphatic portion of $DMAPP^{EA}$, and in turn is a likely factor in determining the relative positions of C3-C^{+} and leaving diphosphate. It is possible the presence or absence of the K258-third conserved region connection influences the Q164-$DMAPP^{EA}$ interaction. If so, not only does the E256D mutation itself potentially influence the relative positions of C3-C^{+} and leaving diphosphate, but also through altering the Q164-$DMAPP^{EA}$ interaction through K258 (Figure 3.7). Regardless, the E256D mutation appears to be a prime determinate in regulating the flux of CPP and LPP formation.

3.8 The Enigmatic F10 Chimera

The F9 to F10 chimeric enzyme transition showed drastic decreases in $K_M^{DMAPP,EA}$ and $K_M^{IPP,ED}$. Comparing CPPase and F10 molecular models do not yield any obvious explanations for the isolated shift in substrate binding. If anything, the F10 chimera is a testament to the dynamic personality of the sagebrush IS-1 fold, where mutations distal from the active site can have consequences altering the catalytic characteristics of the enzymes.

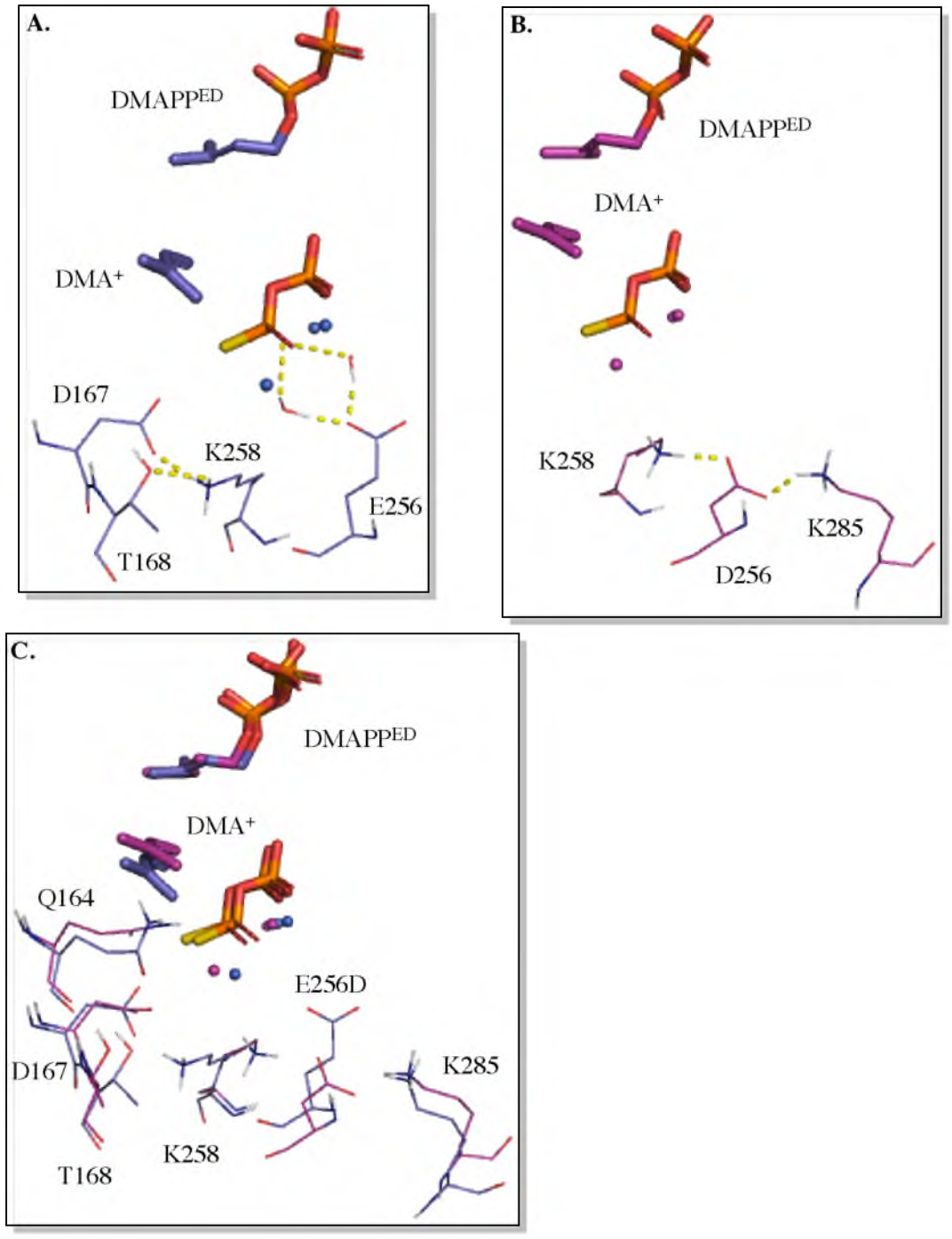


Figure 3.7. Molecular models of CPPase (**A**, blue) and chimera C13 (**B**, magenta) showing the differences in their electron acceptor sites due to the E256D mutation. **C.** Structural overlap of CPPase and C13 molecular models. Magnesium ions are shown.

3.9 The Significance of the C-terminus

Sagebrush FPPase is a well-evolved enzyme designed for the biosynthesis of FPP. It is apparent that sagebrush FPPase provides a genetic template for duplication and mutation, and in doing so produced the catalytically diverse CPPase. Despite the astounding sequential, and therefore structural, similarity of the enzymes, swapping homologous regions between them significantly alters their catalytic properties. None of the chimeric enzymes produced FPP as efficiently as parental FPPase, and similarly, CPPase was the most efficient CPP producer, demonstrating the astonishing capacity of nature to develop and select catalysts.

As discussed above, replacing the C-terminal region of CPPase with FPPase sequence shifted the irregular coupling product ratios, producing an LPPase. Markedly, along the N- to C-terminal conversion of CPPase to FPPase, FPPase-like activity was not restored at any point. In fact, enzymes with C-terminal sequence from CPPase are some of the most catalytically inefficient enzymes studied, accompanying enzymes with the T194G mutation. Clearly the C-terminal region outside of the IS-1 fold is crucial in defining the catalytic nature of a sagebrush enzyme. As exemplified in the CPPase/C13 discussion, notable active site mutations in the C-terminal region are D256E and R343G (FPPase to CPPase). Residue E256 in CPPase participates in binding allylic substrate to the electron acceptor site, whereas R343 in FPPase assists in binding substrate to the electron donor site (Figure 3.8). These C-terminal mutations also have the potential to influence the catalytic efficiency of the chain elongation reaction due to their unique contributions in defining the active site. Regardless of whether or not the two active site mutations are fully responsible for the influence of the C-terminus, the region appears as

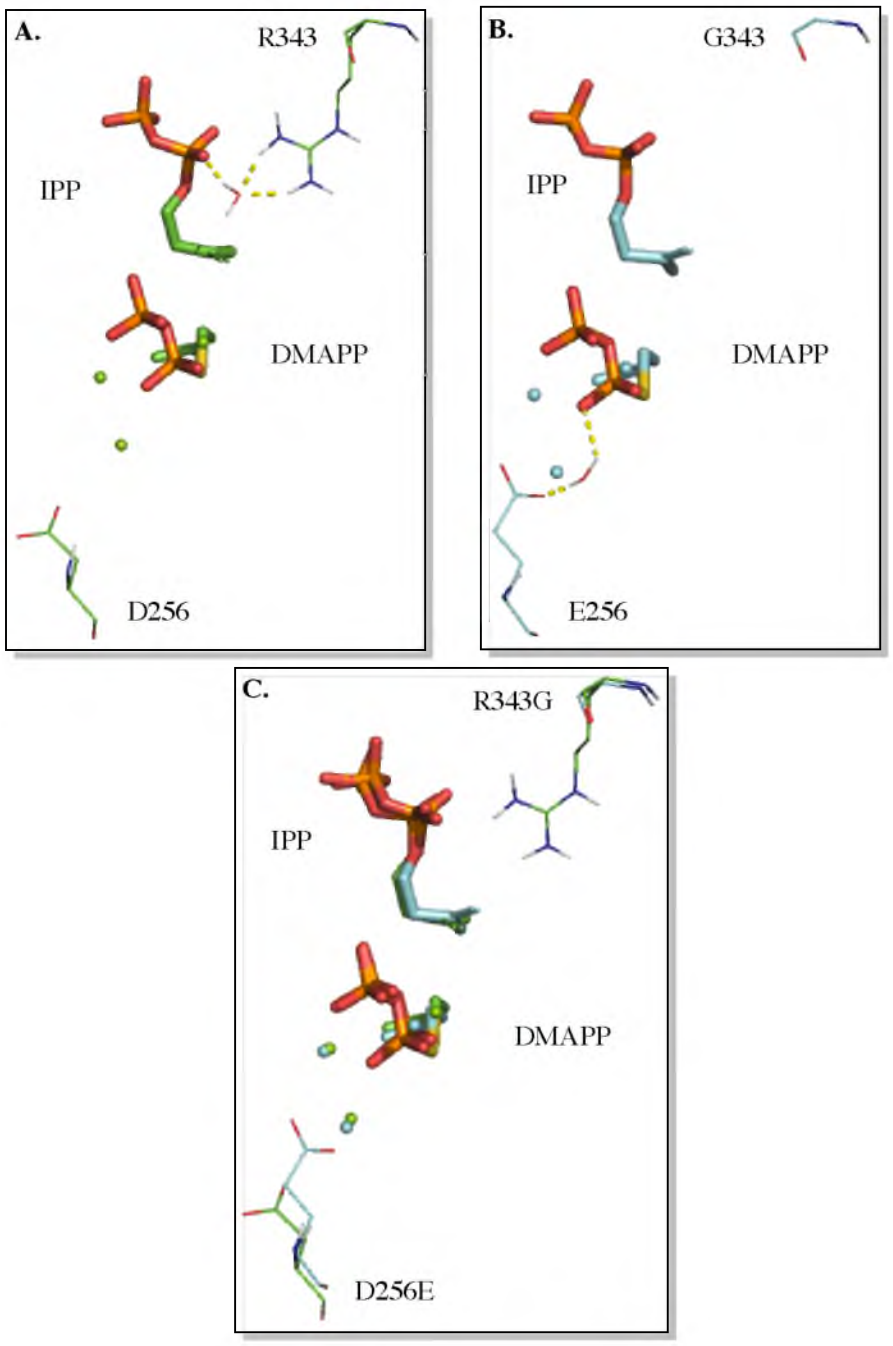


Figure 3.8. Images highlighting the influence of the D256E and R343G FPPase to CPPase C-terminal mutations (**A**, FPPase, green; **B**, CPPase, cyan). **C.** FPPase and CPPase structural overlaps. Magnesium ions are shown.

one of the most significant determinants in defining the catalytic activity and specificity of the sagebrush enzymes.

3.10 References

(1) Tarshis, L. C.; Proteau, P. J.; Kellogg, B. A.; Sacchettini, J. C.; Poulter, C. D. Regulation of Product Chain Length by Isoprenyl Diphosphate Synthases. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 15018-15023.

(2) Stanley-Fernandez, S. M.; Kellogg, B. A.; Poulter, C. D. Farnesyl Diphosphate Synthase. Altering the Catalytic Site To Select for Geranyl Diphosphate Activity. *Biochemistry* **2000**, *39*, 15316-15321.

(3) Thulasiram, H. V.; Erickson, H. K.; Poulter, C. D. A Common Mechanism for Branching, Cyclopropanation, and Cyclobutanation Reactions in the Isoprenoid Biosynthetic Pathway. *J. Am. Chem. Soc.* **2008**, *130,* 1966-1971.

(4) Pyun, H. J.; Coates, R. M.; Wagschal, K. C.; McGeady, P.; Croteau, R. B. Regiospecificity and Isotope Effects Associated with the Methyl-Methylene Eliminations in the Enzyme-Catalyzed Biosynthesis of (R)- and (S)-Limonene. *J. Org. Chem.* **1993**, *58*, 3998-4009.

CHAPTER 4

MATERIALS AND METHODS

4.1 Chimeragenesis

4.1.1 Introduction

A polymerase chain reaction (PCR) method was developed to efficiently build chimeric gene constructs from *Artemisia tridentata* spp. spiciformis (sagebrush) farnesyl diphosphate synthase (FPPase) and chrysanthemyl diphosphate synthase (CPPase) genes. PCR was the method of choice for chimeragenesis, allowing for the selective placement of crossover points and retention of native parental enzyme sequences. Thirteen crossover points were used to build the chimeric gene constructs in regard to the IS-1 fold, with one enzyme turned into the other through sequentially swapping the α-helices and loops of the fold. The chimeric gene construct began with the N-terminal sequence from one parental enzyme up to one of thirteen crossover points, at which point the sequence changes to the other enzyme, and continued as such through the remainder of the construct. To fully assess the influence of the enzymatic structure outside of the IS-1 fold, two additional "core-scaffold" chimeras were built. The additional proteins consisted of the IS-1 fold region from one enzyme (core) with the N- and C-terminal regions of the other.

4.1.2 Cloning of Sagebrush FPPase and CPPase Genes into pET15b

4.1.2.1 Supplementary Cloning Materials and Procedures

4.1.2.1.1 Luria Broth and Agarose Preparation

Luria broth (LB) media was prepared by combining 10 g of tryptone (Fisher-Scientific), 5 g yeast extract (Becton Dickinson Biosciences), and 10 g NaCl (Fisher Scientific) in 1 L of deionized (DI) water. LB agarose was made similarly to LB media, with the addition of 15 g of agar (Fisher Scientific). The LB solution was sterilized by autoclaving. In the case of LB agarose preparation, the autoclaved solution was allowed to cool to room temperature (rt) to 50-55 °C, at which point Ampicillin was added at a concentration of 100 μg/mL (LB+Amp). LB agarose containing Ampicillin was poured into plates and left overnight at rt to solidify, and then stored at 4 °C. After a one-month period, unused LB+Amp agarose plates were discarded. LB solution was stored at rt until used, with Ampicillin added upon use if antibiotic selection was required.

4.1.2.1.2 Gel Electrophoresis Separation of DNA

A 50X stock solution of Tris-acetate-EDTA (TAE) buffer was prepared by adding 242 g of Tris base to 750 mL of DI water. Then 57.1 mL of glacial acetic acid was added, followed by 100 mL of 0.5 M EDTA (pH 8.0). DI water was then used to bring the final volume to 1 L and the solution stored at rt. The final buffer solution had a pH of approximately 8.5.

To prepare the electrophoresis gel, 1.5 g of agarose was added to 150 mL of TAE buffer. The mixture was microwaved to boiling, and then mixed to completely dissolve the agarose. Before casting the TAE agarose solution, 1.5 μL of 10,000X SYBR Safe

DNA Gel Stain (Life Technology) was added in order to visualize the DNA bands. The DNA fragments were separated by applying 125 V in TAE buffer until the orange dye band (approximately 50 bp) from the 6X Blue/Orange DNA Loading Dye (Promega) reached the bottom of the gel. A GeneRuler 1 kb DNA Ladder (Thermo Scientific) was used to estimate the basepair length of sample DNA bands, with the DNA fragments being visualized using a Dark Reader Transilluminator (Clare Chemical).

4.1.2.1.3 Preparation of Chemically Competent DH5α Cells

All materials (centrifuge tubes, bottles, solutions, etc.) used in the preparation of the cells were sterilized by autoclave. A frozen glycerol stock of DH5α cells was used to inoculate 1 mL of LB media, and the sample shaken at 225 rpm and 37 °C overnight. The entire 1 mL culture was then used to inoculate 500 mL of LB, and the sample shaken at 225 rpm and 37 °C until the optical density at 550 nm (OD_{550}) reached 0.45-0.6 (~ 4 h). The 500 mL culture was then transferred in 250 mL portions to two 500 mL centrifuge tubes and chilled to 4 °C. The samples were then centrifuged at 6,000 g for 5 min at 4 °C, and then the supernatant removed. Each cell pellet was then resuspended in 125 mL of ice-cold 50 mM $CaCl_2$, and the samples combined into a single 500 mL centrifuge bottle. The sample was then centrifuged at 6,000 g for 5 min at 4 °C. After centrifugation, the supernatant was poured off and then the cell pellet was resuspended in 21.5 mL of ice-cold 50 mM $CaCl_2$. After cell resuspension, 3.5 mL of glycerol was added to the sample and the solution gently mixed. The cell suspension was aliquotted into sterile 1.7 mL microcentrifuge tubes in 500 μL portions, flash frozen using liquid nitrogen, and then stored in a -80 °C freezer until use.

4.1.2.1.4 Transformation of DNA into Chemically Competent DH5α Cells

A frozen glycerol DH5α cell stock was thawed on ice. After thawing, a 100 μL portion(s) of the cell stick was transferred to 1.7 mL microcentrifuge tubes on ice. A 1 μL portion of plasmid DNA was added to the cells and the sample incubated on ice for 20-25 min. The sample was then heat shocked by placement in a 42 °C water bath for 45 s, and then promptly returned to ice. A 1 mL portion of Super Optimal broth with Catabolite repression (SOC medium, Invitrogen) was added, and after which the sample transferred to a 14 mL polystyrene or polypropylene Falcon round-bottom tube(s) (Fisher Scientific) and shaken at 225 rpm and 37 °C for 1 h. Aliquots of 2, 20, and 200 μL were then plated on LB+Amp agarose and the plates incubated overnight at 37 °C.

4.1.2.1.5 Restriction Endonuclease Digestion of Plasmid DNA

Plasmid DNA used for restriction enzyme-based cloning was always "doubly digested," which is digestion by two restriction enzymes at a time. The restriction endonucleases used to clone the sagebrush genes into pET15b were FastDigest NdeI, BamHI, and Mva1269I (Thermo Scientfic). A 1 μL portion of each enzyme was used to doubly digest 10 μg of plasmid DNA in a total volume of 20 μL at 37 °C for 1 h. Thermo Scientific's proprietary FastDigest Green Buffer was used in each reaction.

4.1.2.1.6 Miniprep and Midiprep Purification of Plasmid DNA from Cell Cultures

To prepare a cell culture for the miniprep purification of plasmid DNA, a single cell colony was used to inoculate 5 mL of LB+Amp media. The inoculated media was

then shaken at 225 rpm and 37 °C overnight. The plasmid DNA was purified from the 5 mL overnight culture using a QuickClean 5M Miniprep Kit (GenScript).

To prepare a cell culture for midiprep purification, 1 mL of LB+Amp was inoculated, as described above, and shaken in a water bath at 225 rpm and 37 °C for 6-8 h. Following the incubation, 50 μL of the preculture was used to inoculate 50 mL of LB+Amp, and the 50 mL inoculation was shaken at 225 rpm and 37 °C overnight. The plasmid DNA was then purified from the 50 mL culture using a CompactPrep Plasmid Midi Kit (Qiagen).

4.1.2.1.7 Purification of DNA Bands from Agarose

DNA was purified from polymerase chain reactions and gel electrophoresis experiments using an illustra GFX PCR DNA and Gel Band Purification Kit (GE Life Science).

4.1.2.1.8 Transformation of Plasmid DNA Into XL10 Gold Ultracompetent Cells

After thawing on ice, a 30 μL portion of XL10 Gold ultracompetent cells (Stratagene) was pipetted into an ice-cold 1.7 mL microfuge tube. Then a 1.2 μL aliquot of the provided β-mercaptoethanol (β-Me) solution was added to the sample and the tube swirled gently. The sample was then incubated on ice for 10 min, with gentle swirling every 2 min. A 3 μL portion of plasmid DNA was then added to the cells, the tube gently swirled, and the sample incubated on ice for 30 min. The sample was then heat-pulsed by placement in a 42 °C water bath for 30 s, and then promptly removed and incubated on

ice for 2 min. A 200 μL portion of SOC medium was then added and the sample shaken in a water bath at 225 rpm and 37 °C for 1 h. After shaking, a 2-200 μL portion of the cell culture was pipetted onto LB+Amp agarose plates and incubated overnight at 37 °C.

4.1.2.2 Preparation of NdeI and BamHI Restriction Enzyme Digestions of Sage FPPase and CPPase Genes

The DNA sequences for sagebrush FPPase and CPPase genes were synthesized in *E. coli* context by GenScript and obtained as NdeI/BamHI inserts in the plasmid pUC57. The pUC57-FPPase and pUC57-CPPase plasmids were chemically transformed into DH5α cells. Then minprep purifications were performed on the pUC57-gene-DH5α samples to obtain working stock solutions of pUC57-FPPase and pUC57-CPPase. NdeI and BamHI restriction enzyme digestions were performed on 10 μg of pUC57-FPPase and pUC57-CPPase, and the DNA fragments were separated by gel electrophoresis. DNA bands corresponding to the genes (~ 1375 base pairs (bp)) were excised from the gel, purified, and stored at -20 °C until needed.

4.1.2.3 Preparation of NdeI/Mva1269I and BamHI/Mva1269I Digested pET15b

It proved difficult to purify NdeI/BamHI digested pET15b from undigested plasmid, as the excised DNA fragment produced by the digestion is only 12 bp long. In order to separate NdeI- and BamHI-digested pET15b, a third restriction enzyme was used. A 10 μg portion of pET15b was submitted to NdeI/Mva1269I and BamHI/Mva1269I double digestions. The DNA gel electrophoresis band showing a

length of approximately 3017 bp from the NdeI/Mva1269I digestion, and the band showing a length of approximately 2391 bp from the BamHI/Mva12691I, were gel purified and stored at -20 °C until needed.

4.1.2.4 Production of pET15b Containing Sagebrush FPPase and CPPase NdeI/BamHI Inserts

New England Biolabs's (NEB) Quick Ligation Kit was used for the 3-piece ligation. A 3 μL portion (each) of the NdeI/BamHI digested gene (obtained from the pUC57-gene digestions), the DNA fragment purified from the NdeI/Mva12691I digestion of pET15b, and the DNA fragment purified from the BamHI/Mva12691I digestion of pET15b were combined in a 1.7 mL microfuge tube. A 10 μL portion of the 2X Quick Ligation Buffer and 1 μL portion of the Quick T4 DNA Ligase was added and the solution thoroughly mixed. The sample was then briefly centrifuged and incubated at rt for 30 min. After the ligation reacton, the sample was transformed into XL10 Gold ultracompetent cells, plated on LB+Amp agarose, and incubated overnight at 37 °C. A miniprep was then performed on 5 colonies and the purified plasmid DNA screened for ligation product via gel electrophoresis – a band showing a length of approximately 7083 bp indicated the formation of the pET15b-gene product. Sanger DNA sequencing by the University of Utah DNA Sequencing core, using universal T7 forward and reverse primers, verified pET15b-FPPase and pET15b-CPPase constructs.

4.1.3 Construction of Chimeric Genes Using the Polymerase Chain Reaction

4.1.3.1 Introduction

The method used to construct the chimeric protein genes was based upon megaprimer whole plasmid (MEGAWHOP) and structure-based combinatorial protein engineering (SCOPE) polymerase chain reaction (PCR) techniques.[1,2] The genes were assembled through two rounds of PCR using PfuUltra II Fusion HS DNA polymerase (Agilent) and an MJ Research PTC-200 PCR thermal cycler. The first reaction used a forward primer and a "crossover" reverse primer to construct a megaprimer, which contains the coding region for the N-terminal sequence of the chimeric protein. The second reaction completes the crossover from one enzyme's DNA sequence to the other, building the remaining C-terminal sequence onto the megaprimer and finalizing the chimeric gene construct in pET15b.

All primers were designed using Integrated DNA Technologies' OligoAnalyzer to calculate their melting temperatures and synthesized by University of Utah's DNA/Peptide Synthesis core. In the case of the crossover primers, each half was designed with a 60 °C melting temperature. The forward primers were designed from N-terminal CPPase and FPPase to have a 60 °C melting point, with a 20 base pair sequence of complementary pET15b sequence added to the 5' end to assist megaprimer binding in the second PCR reaction.

4.1.3.2 Polymerase Chain Reactions Employed in the Construction of Chimeric Genes

For the first reaction, 50 ng of vector template (pET15b-gene), 250 μM deoxynucleotides (dNTPs), 200 nM of the forward primer (dependent upon whether FPPase or CPPase is the template gene), 200 nM of the reverse "crossover" primer, 10 μL of the proprietary PfuUltra II 10X reaction buffer, and 2 μL of PfuUltra II were combined in a total volume of 100 μL. The samples were then heated to 95°C for 2 min, followed by 30 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 21 s, with an additional final extension at 72°C for 3 min. The sample was then cooled to 4 °C, and the megaprimer purified by gel electrophoresis.

For the second reaction, 120 ng of template DNA (appropriate pET15b-gene), 250 μM dNTPs, 5 μL of PfuUltra II 10X reaction buffer, 1 μL of PfuUltra II, and the appropriate amount of purified megaprimer to bring the total volume to 50 μL, were combined. The thermocycle program for the second reaction was the same as the first, except the time for the 72 °C primer extension phase was extended to 1 min 47 s to allow the complementary strand for the entire plasmid to be synthesized.

To cut the template DNA, the PCR reaction solution was submitted to FastDigest DpnI (Fermentas) digestion by adding 7 μL of the FastDigest Green Buffer, 7 μL of DpnI, and 6 μL of DI water. The reaction solution was mixed and incubated at 37 °C for 1 h. Following DpnI digestion, 3 μL of the mixture was transformed into 30 μL of XL10 Gold cells, with 20 and 200 μL portions plated on LB+Amp agarose plates. After incubating the plates overnight at 37 °C, minipreps were performed on 5 colonies. The purified plasmid DNA was screened via gel electrophoresis, and samples showing proper

size (~ 7083 bp) were screened by Sanger DNA sequencing. Samples showing production of the desired chimeric gene were stored at -80 °C (Figure 4.1).

4.2 Expression and Purification of Enzymes

4.2.1 Supplementary Materials and Methods

4.2.1.1 Stock Solutions

Deionized water was used in the preparation of all solutions. After preparation, the solutions were sterilized by autoclave, unless indicated otherwise.

4.2.1.1.1 Preparation of 50X 5052 Solution

A 1X 5052 solution contains 0.5% glycerol, 0.05% glucose, and 0.2% α-lactose. To make 100 mL of 50X 5052 solution, 25 g of glycerol was added to 73 mL of water and the solution mixed. Then, a 2.5 g portion of glucose and a 10 g portion of α-lactose monohydrate were added and the solution mixed. The solution was then heated by microwave to facilitate dissolution.

4.2.1.1.2 Preparation of ZY Broth

Adding 10 g of tryptone and 5 g of yeast extract to 1 L of water made ZY broth.

4.2.1.1.3 Preparation of 25% Aspartate Solution (w/v) Solution

Combining 25.0 g (0.188 mol) of aspartic acid and 8.0 g (0.20 mol) of NaOH in 84 mL of water made a 25% aspartate solution.

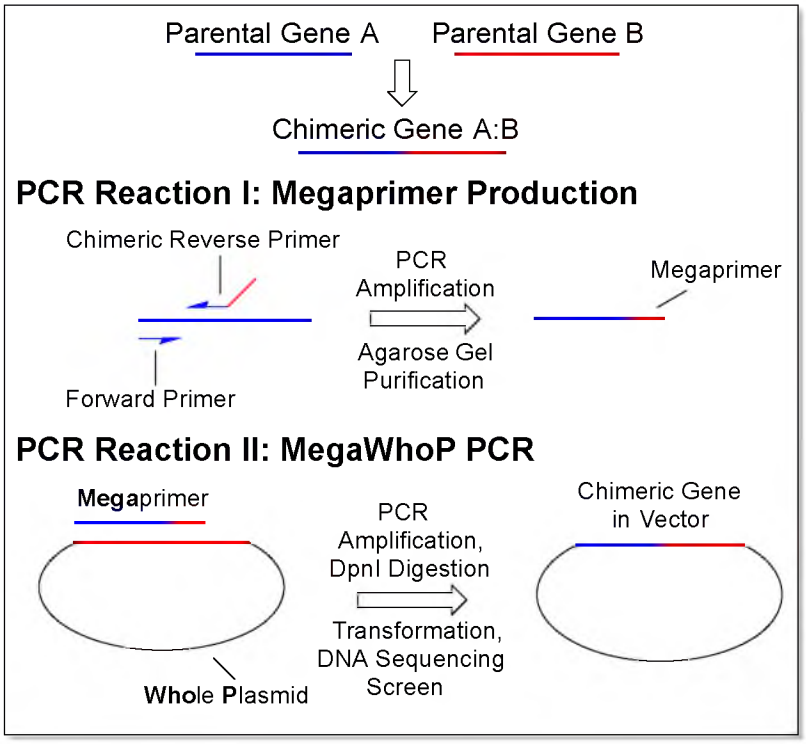


Figure 4.1. Summary of the PCR reactions used for chimeragenesis. The first reaction produces the megaprimer, which is used to finalize building a chimeric enzyme construct via megaprimer whole plasmid PCR (MegaWhoP PCR).

4.2.1.1.4 Preparation of 50X M Solution

A 1X M solution contains 50 mM PO_4^{-3}, 50 mM NH_4Cl, and 5 mM Na_2SO_4. Combining 17.75 g of Na_2HPO_4, 17.0 g of KH_2PO_4, 13.4 g of NH_4Cl, and 3.55 g of Na_2SO_4 in 80 mL of water produces a 100 mL portion of the 50X stock solution. The 50X solution may form crystals over time that redissolve upon heating.

4.2.1.1.5 Preparation of 0.1 M $FeCl_3$ in ~0.12 M HCl Solution

The $FeCl_3$ in HCl solution was not autoclaved, as a large precipitate forms. A 1 mL portion of concentrated HCl (approximately 12 M) and 2.7 g of $FeCl_3$-$6H_2O$ was added to 99 mL of sterile water to produce the 0.1 M $FeCl_3$ solution.

4.2.1.1.6 Preparation of the 1000X Trace Metal Solution

The following stock solutions of individual metals were prepared and autoclaved: 1 M $CaCl_2$, 1 M $MnCl_2$-$4H_2O$, 1 M $ZnSO_4$-$7H_2O$, 0.2 M $CoCl_2$-$6H_2O$, 0.1 M $CuCl_2$-$2H_2O$, 0.2 M $NiCl_2$-$6H_2O$, 0.1 M Na_2MoO_4-$2H_2O$, 0.1 M Na_2SeO_3-$5H_2O$, and 0.1 M H_3BO_3. To produce 100 mL of the 1000 X trace metal solution 50 mL of 0.1 M $FeCl_3$ in ~0.12 M HCl, 2 mL of 1 M $CaCl_2$, 1 mL of 1 M $MnCl_2$-$4H_2O$, 1 mL of 1 M $ZnSO_4$-$7H_2O$, 1 mL of 0.2 M $CoCl_2$-$6H_2O$, 2 mL of 0.1 M $CuCl_2$-$2H_2O$, 1 mL of 0.2 M $NiCl_2$-$6H_2O$, 2 mL of 0.1 M Na_2MoO_4-$2H_2O$, 2 mL of 0.1 M Na_2SeO_3-$5H_2O$, 2 mL of 0.1 M H_3BO_3, and 36 mL of sterile water were combined. The final 1000X trace metal solution was not autoclaved.

4.2.1.1.7 Preparation of the 18 Amino Acid Solution

A solution containing 1 g of the following 17 amino acids (aa) in 90 mL of water was produced: sodium glutamate, aspartic acid, lysine, arginine, histidine, alanine, proline, glycine, threonine, serine, glutamine, asparagine, valine, leucine, isoleucine, phenylalanine, and tryptophan. The "17 aa" solution was then passed through a 0.2 um filter (not autoclaved) and stored at 4 °C.

The 18^{th} amino acid is methionine and was stored separately as a 25 mg/mL solution. The working 18 amino acid solution (18 aa) was produced by mixing the two solutions in a 5:2::17 aa:methionine (v/v) ratio.

4.2.1.2 Auto-Induction Medias

4.2.1.2.1 Preparation of MDG Media

MDG is a noninducing, minimal media. A 10 mL portion of MDG is produced by combining 20 μL of 1 M $MgSO_4$, 2 μL of 1000X trace metals, 125 μL of 40% glucose (w/v), 100 μL of 25% aspartate (w/v), and 200 μL of 50X M in 9.55 mL of water.

4.2.1.2.2 Preparation of MDAG-11 Agarose

MDAG-11 is a noninducing, minimal medium supplemented with amino acids. To make 500 mL of agar, 5 g of agarose was added to approximately 475 mL of water, and then the solution was autoclaved. Upon removal from the autoclave, the solution was mixed well and allowed to cool on the bench top for approximately 10 min. Then 1 mL of 1 M $MgSO_4$, 100 μL of 1000X trace metals, 1.25 mL of 40% glucose (w/v), 2 mL of 25% aspartate (w/v), 10 mL of 50 X M, 14 mL of 18 aa, and ampicillin to a final concentration of 100 μg/mL was added. The solution was then mixed and the agarose plates poured.

4.2.1.2.3 Preparation of ZYM-5052 Media

ZYM-5052 is a complex auto-inducing medium. A 10 mL portion of ZYM-5052 was made by adding 20 μL of 1 M $MgSO_4$, 2 μL of 1000X trace metals, 200 μL of 50X 5052, and 200 μL of 50X M to 9.57 mL of ZY broth.

4.2.1.3 Chemical Transformation of Plasmid DNA into BL21 Star (DE3) Cells

A vial containing 50 μL of BL21 Star (DE3) cells was thawed on ice. A 1 μL portion of pET15b-gene was added to the cells and the sample gently mixed by tapping. The sample was then incubated on ice for 30 min. After incubation on ice, the sample was heat-shocked by placement in a water bath at 42 °C for 30 s, after which the cells were promptly removed and placed back on ice. A 250 μL portion of SOC media was added to the sample, and then shaken at 225 rpm and 37 °C for 1 h. Afterwards, 20 and 200 μL portions were plated on LB+Amp agarose and the plates incubated overnight at 37 °C.

4.2.1.4 Production of BL21 Star (DE3) Expression Cell Stocks

Plasmids with genes for the chimeric proteins inserted into pET15b were chemically transformed into BL21 Star (DE3) cells, plated on MDAG-11 plates containing 50 μg/mL ampicillin, and incubated overnight at 37 °C. Two colonies were used to inoculate 2 mL of MDG media containing 50 μg/mL ampicillin. The samples were shaken overnight at 225 rpm and 37 °C. Glycerol (100 μL) was added to a 1 mL portion of each culture, the samples were lightly vortexed, and then stored at -80 °C until needed.

4.2.1.5 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

A 1-5 μL portion of protein-containing sample and 5 μL of 4X NuPAGE LDS Sample Buffer (Life Technologies) were combined and the total volume of the mixture brought to 20 μL. The sample tube was incubated in boiling water for 10-15 min. After boiling, the sample was cooled at room temperature for 10-15 min. The protein sample was then loaded onto a NuPAGE Tris-Acetate Gel (Life Technologies) and the gel developed by applying 150 V in 3-(N-morpholino)propanesulfonic acid (MOPS) buffer (50 mM MOPS, 50 mM Tris base, 0.1% SDS, 1 mM EDTA, pH 7.7). The protein bands were visualized by staining with GelCode Blue Safe Protein Stain (Thermo Scientific), and the molecular weights of the protein bands were estimated by comparison with a PageRuler Plus Prestained Protein Ladder (Thermo Scientific, 10-250 kDa range).

4.2.2 Expression of Proteins

4.2.2.1 Introduction

Studier-based auto-induction was used for protein expression.[3] The auto-induction procedure gave an approximately 10-fold increase in soluble sagebrush FPPase over induction with isopropyl β-D-thiogalactoside (IPTG). Expression was further enhanced using BL21 (DE3) Star cells (Life Technology), which have a truncation that deletes the C-terminal domain of the RNase E enzyme, preventing the degradation of mRNA.

4.2.2.2 Auto-induction of Expression

A sterile pipette tip was scraped across the top of a pET15b-gene-BL21 (DE3) Star frozen cell stock and used to inoculate 2 mL of MDG media, the inoculation culture was then developed by shaking at 225 rpm overnight at 37 °C. A 400 µL portion (each) of the preculture was then used to inoculate two flasks containing 400 mL of ZYM-5052 containing 50 µg/mL of ampicillin. The samples were shaken at 300 rpm and 37 °C for 6-8 h, and then allowed to cool to room temperature. Shaking continued for another 14-16 h (22-24 h total). Autoinduction of expression occurs as the culture's cell growth reaches saturation. Sage FPPase and CPPase typically showed saturation of cell growth at an optical density at 600 nm (OD_{600}) between 6-8. After shaking, the samples were cooled to 4 °C, and the cells harvested by centrifugation at 4000 g and 4 °C. The 400 mL cultures were combined and centrifuged into a single cell pellet, and the supernatant poured off. The pellets were stored at -80 °C until needed.

4.2.3 Purification and Storage of Proteins

Protein purification was performed at 4 °C. The cell pellets were removed from the freezer and thawed, followed by addition of 20 mL of lysis buffer (20 mM sodium phosphate at pH 7.4, 500 mM sodium chloride, 20 mM imidazole, 10 mM β-mercaptoethanol), Tween 20 to 0.1% (v/v), DNase I (5 µg/mL), RNase A (10 µg/mL), 2 Roche Mini Complete protease inhibitor cocktail tablets (EDTA-free), and 20 mg of lysozyme. The cell pellet was fully suspended and the sample shaken at 120 rpm for 30 min. The samples were then sonicated on ice 6 times for 10 s each (70% duty cycle and output control at 4), with a minimum 1-min cooling period between each sonication. The

samples were then centrifuged at 10,000 g for 30 min, and the supernatant passed through a 0.2 μm filter.

HisTrap HP-5 columns (GE Life Science) were equilibrated with lysis buffer and the filtered supernatants were loaded onto the column via peristaltic pump at a flow rate of ~2 mL/min. The HisTrap column was then transferred to a GE Life Science AKTA FPLC system, where lysis buffer was passed through the column at a flow rate of 3 mL/min until the ultraviolet (UV) absorbance at 280 nm stabilized, indicating the removal of unbound protein. Column-bound protein was eluted using a linear gradient of 0 to 100% elution buffer (500 mM imidazole, 20 mM sodium phosphate buffer, pH 7.4, 500 mM sodium chloride, 10 mM β-mercaptoethanol) over a period of 120 min. Five-mL fractions were collected during protein elution, and fractions containing protein (as indicated by UV absorbance at 280 nm) were analyzed by SDS-PAGE. The protein fractions containing highly pure protein (> 95%) were combined. In cases where protein purification levels were low, every other fraction in the expected range of elution was analyzed by SDS-PAGE, and the fractions containing chimeric protein were combined. The purified proteins were then precipitated by adding 470 mg of ammonium sulfate added to each mL of solution. The protein-ammonium sulfate solutions were allowed to stand overnight at 4 °C to facilitate precipitation.

The suspension of precipitated protein was centrifuged at 7000 g for 30 min, the supernatant removed, and the protein precipitate dissolved in a minimal amount of lysis buffer. The purified protein solution was then transferred to a Slide-A-Lyzer dialysis cassette (Pierce, 10K MWCO) and dialyzed twice against 20 mM Tris buffer, pH 8, with 500 mM sodium chloride, and then once against the same solution also containing 20%

glycerol. Each dialysis was performed for a minimum of 4 h at a minimum ratio of 1:100 (v/v) protein solution:buffer.

Flash frozen droplets of the protein solution were produced by filling a 50 mL plastic centrifuge tube with liquid nitrogen, immersing the tube in liquid nitrogen, and then slowly adding dialyzed protein solution drop by drop into the tube. After which the container was removed from the liquid nitrogen, loosely capped, and placed in the -80 °C freezer. Once the liquid nitrogen evaporated, the container cap was tightened and the pellets stored at -80 °C. This protocol conveniently produced approximately 100 μL, frozen spherical samples of protein. Enzyme concentrations were determined using the BCA Protein Assay Kit (Pierce). All enzymes except those that did not express well (C9, F2-F7) were desalted with 0.1% formic acid via Amicon Ultra 0.5 mL centrifugal filters (EMD Millipore, 10K molecular weight cut-off (MWCO)) and submitted to the mass spectrometry facility for molecular weight verification by electrospray ionization mass spectrometry (ESI-MS) - many proteins precipitated upon desalting and could not be processed.

4.3 Synthesis of DMAPP, IPP, and GPP

Diphosphate substrates were synthesized from the corresponding alcohols according to the procedure of Davisson and coworkers and stored as solutions in 50 mM ammonium bicarbonate at -80 °C.[4] Concentrations of the diphosphates were determined by a malachite green-based assay.[5]

4.4 Enzymatic Product Ratio Studies

FPPase, CPPase, and chimeras of the two proteins were incubated with diphosphate substrates in 35 mM HEPES buffer, pH 7.6, containing 10 mM magnesium chloride, and 10 mM β-mercaptoethanol (reaction buffer), with only minor variations from the protocols described by Thulasiram and coworkers.[6]

4.4.1 Chain Elongation Products

Enzymatic products formed under chain elongation conditions were obtained by combining 100-150 μg of purified enzyme (or as much as possible in the cases of the F2-F7 chimeras due to poor soluble enzyme expression), 500 μM IPP, 500 μM DMAPP, and 20 μL of 10X reaction buffer in a total volume of 200 μL. Each reaction was performed in triplicate. The samples were then incubated at 30 °C for 2 h. The diphosphate moieties were removed by the addition of 12 μL of NEB Buffer 3 (1 M NaCl, 500 mM Tris-HCl, 100 mM $MgCl_2$, 10 mM dithiothreitol, pH 7.9) and 80 units of calf intestinal phosphatase (NEB), followed by incubation at 37 °C for 1 h. The samples were transferred to 100 X 13 mm glass test tubes, and the reaction vessel washed with 600 μL of water and then also added to the test tube. A 1 g portion of sodium chloride was added, and the sample was extracted 3 times with 800 μL of methyl tert-butyl ether (MTBE). The MTBE extract was concentrated under a gentle stream of nitrogen until the volume was reduced to ~ 75 μL. The extract was then transferred to a GC vial containing a 150 μL glass insert (Agilent) and sealed with a crimp top.

The concentrated MTBE extract was analyzed by gas chromatography with flame ionization detection (GC-FID) analysis using the following protocol:

- Agilent HP-5 capillary column (30 m X 0.320 mm, 0.25 micron)
- Ultra Inert Inlet Liner (Agilent, splitless, single taper with wool)
- He flow rate of 1 mL/min
- 5 μL splitless injection by autosampler
- Inlet at 230 °C, detector at 250 °C
- Temperature program: 60 °C to 120 °C at 2 °C/min; 120 °C to 230 °C at 10 °C/min; hold at 230 °C for 10 min.

Product ratios were determined using peak integration software (Agilent's ChemStation), with an assumed signal response factor of one. The structures of the products were established by comparison of retention times and mass spectra with authentic samples.[6]

4.4.2 Irregular Products

Products formed under irregular terpenoid coupling conditions were obtained by combining 100-150 μg of purified enzyme (or as much as possible in the cases of the F2-F7 chimeras due to poor soluble enzyme expression), 3 mM DMAPP, and 20 μL of 10X reaction buffer in a total volume of 200 μL. Each enzyme was analyzed in triplicate. The sample was incubated at 30 °C for ~ 16 h, after which concentrated MTBE extracts were prepared for GC analysis as described above. The concentrated MTBE extracts were analyzed by GC-FID using the following protocol:

- Agilent HP-Chiral capillary column (30 m X 0.320 mm, 0.25 micron)
- Ultra Inert Inlet Liner (Agilent, splitless, single taper with wool)
- He flow rate of 1 mL/min
- 5 μL splitless injection by autosampler

- Inlet at 230 °C, detector at 250 °C
- Temperature program: hold at 50 °C for 3 min; 50 °C to 120 °C at 2 °C/min; 120 °C to 230 °C at 10 °C/min; hold at 230 °C for 10 min

Product ratios were determined as described above, with the structures of products were determined by comparison of retention times and mass spectra with authentic samples.[6]

4.5 Determination of Apparent Michaelis-Menten Kinetic Parameters

4.5.1 Synthesis of [β-^{32}P]DMAPP and [β-^{32}P]IPP

[β-^{32}P] Labeled diphosphate substrates were prepared in two steps as described by Chen and coworkers.[7] Dimethylallyl phosphate (DMAP) and isopentenyl phosphate (IP) were synthesized from their corresponding alcohols, and the [β-^{32}P] labeled diphosphate substrates were prepared from their corresponding monophosphates. In a typical procedure, 65 nmol of monophosphate was incubated with 13 nmol of [γ-^{32}P]ATP (10 Ci/mmol, Perkin Elmer), ~ 10 ng of *Thermoplasma acididophilum* (THA) isopentenyl phosphate kinase (IPK) in 100 μL of IPK reaction buffer (100 mM HEPES buffer, pH 7.5, 10 mM $MgCl_2$, 10 mM β-mercaptoethanol, and 1 mg/mL bovine serum albumin (BSA)) at 30 °C overnight. The radiolabeled products were purified by cellulose column chromatography.[8] Fractions were collected, and those containing substantial radioactivity, as determined with a Geiger counter, were combined in a round bottom flask. Solvent was removed to a volume of approximately 1 mL by rotary evaporation and the samples were stored at -20 °C until used.

4.5.2 Kinetic Experiments

Kinetic assays were performed in 35 mM HEPES buffer, pH 7.6, containing 10 mM $MgCl_2$, 10 mM β-mercaptoethanol, and 2 mg/mL of BSA in a total volume of 50 μL at 30 °C. For chain elongation, unlabeled IPP was spiked with 1 μL of [^{32}P]IPP, and for formation of irregular terpenes, unlabeled DMAPP was spiked with 1 μL of [^{32}P]DMAPP. Samples were pre-incubated for 5 min at 30 °C before the reactions were initiated by the addition of enzyme. Typically, samples were incubated for 10 min, although it was necessary to increase the time to 20 min when turnover was slow. Assay mixtures were quenched with 125 μL of a (100:13, v/v) solution methanol and 750 mM EDTA. A 20 μL portion of the quenched sample was spotted onto a silica TLC plate, and the plates were developed with $CHCl_3$/pyridine/formic acid/H_2O (30:70:16:10, v/v/v/v). The TLC plates were imaged using a storage phosphor screen and scanning with a Typhoon 8600, with the radiological counts quantified using ImageQuant 5.2. Single point assays were measured in triplicate under initial velocity conditions where turnover of substrates was < 10%, with an upper limit of substrate concentrations at 6 mM.

4.5.2.1 Chain Elongation – IPP and DMAPP

A series of experiments were performed to deduce the necessary boundary conditions for substrate and enzyme concentrations. Enzyme-free experimental controls did not show detectable background for formation of chain elongation products. Preliminary screens used 1X, 10X, 100X, and 1000X dilutions of purified enzyme and 74 μM, 222 μM, 667 μM, 2000 μM, and 6000 μM concentrations of both DMAPP and IPP. The lowest combination [DMAPP] and [IPP] concentrations (the "optimal"

concentrations) needed to give a maximal velocity was determined, and the [enzyme] required for ~ 1% turnover at these concentrations of IPP and DMAPP was estimated.

Then the initial rates were determined for varied [IPP] using the "optimal" [DMAPP] and [enzyme] determined above. The [IPP] concentration observed to maximize the initial rate became the "new optimum" [IPP]. A second set of assays was performed with varied [DMAPP] using an [enzyme] estimated to give 5% turnover based on the "new optimum" [IPP]. A "new optimum" [DMAPP] was determined, and the enzyme concentrations producing approximately 1% and 5% turnovers at maximized k_{cat} calculated.

Apparent values for Michaelis-Menten constants for IPP and DMAPP were obtained by varying the concentrations of one substrate, while holding the concentration of the other substrate at the "new optimum" value. The [enzyme] giving ~1% turnover was used for the varying [IPP] experiment, whereas the [enzyme] giving ~5% turnover was used when varying [DMAPP].

4.5.2.2 Chain Elongation – GPP

Apparent values of Michaelis-Menten constants for GPP were obtained as described above using the "new optimum" [IPP].

4.5.2.3 Irregular Terpenes – DMAPP

Assays were typically conducted using 2 μL of undiluted purified enzyme stock solutions and varied [DMAPP] between 74 μM to 6000 μM over 20 min at 30 °C. Enzyme-free experimental controls confirmed background formation of irregular

terpenes, which were accounted for in the calculation of apparent Michaelis-Menten values.

4.6 Construction of Molecular Models

Molecular models of substrate-bound enzymes were built from an apo C3 chimeric enzyme X-ray crystal structure (Brookhaven pdb 4KK2) and a DMASPP- and IPP-bound *E. coli* FPPase x-ray crystal structure (Brookhaven pdb 1RQI).[9] The software programs PyMol and UCSF Chimera were used to develop the models.[10] First, all non-active site water molecules were deleted from structure 1RQI. The residues G174 and E175 were missing from one monomer of structure 4KK2, and the missing residues were modeled into the crystal structure using UCSF Chimera's Model/Refine Loops tool.[11] The 1RQI structure was opened in UCSF Chimera and the refined 4KK2 structure overlaid on 1RQI. Structure 4KK2 is in the substrateless "open" conformation, whereas structure 1RQI is "closed." The 4KK2 structure was taken to the closed conformation (4KK2_morph) by morphing the structure into 1RQI via the Morph Conformations tool. UCSF Chimera's DockPrep tool can perform several tasks including deleting water molecules, repairing truncated side chains, adding hydrogens, assigning partial charges, and writing files in mol2 format. The tool is mainly utilized here to add hydrogens after adding or mutating residues, and to assign partial charges. At this point a DockPrep was performed on structure 4KK2_morph and a mol2 file saved (4KK2_morph_DockPrep).[12]

Structure 4KK2_morph_DockPrep was then overlapped with 1RQI in UCSF Chimera. The 1RQI sequence was opened, selected, and deleted, leaving DMASPP, IPP, and the active site waters behind in the 4KK2 structure. The Copy/Combine Molecular

Models tool was used to merge the two structural files into one (4KK2_overlap) and a DockPrep performed. Structure 4KK2_overlap was the parental structure from which all others models were built.

To build model structures of FPPase, CPPase, and select chimeric enzymes, amino acid mutations were performed on 4KK2_overlap using UCSF Chimera's Rotamers tool. The highest-probability Dunbrack rotamer was selected for a mutation, except in the case of F231Y and D235N FPPase to CPPase mutations, where the rotomers best matching the amino acid conformations in structure 4KK2 were chosen.[13] A DockPrep was then performed on the model, followed by optimization of the energetic parameters of the models over 2000 steps (1000 steepest descent and 1000 conjugate gradient).[14]

To build model structures with DMAPP in the electron donor site, IPP's valency was cycled in Pymol, isomerizing the double bond and turning the substrate into DMAPP before performing the final DockPrep and energy minimization. In the production of models with the dimethyallyl carbocation (DMA+) and diphosphate in the electron acceptor site, the dimethyallyl (DMA)-diphosphate bond was removed in Pymol, and then bonds made between DMAPP in the electron donor site and DMA in the electron acceptor site to produce chrysanthemyl diphosphate (CPP) before performing the final DockPrep and energy minimization. The bonds formed to make CPP in the active site were then removed after energy minimization, and the appropriate valencies assigned, leaving DMAPP in the electron donor site and DMA+ and diphosphate in the electron acceptor site.

4.7 References

(1) Miyazaki, T.; Takenouchi, K. Creating Random Mutagenesis Libraries Using Megaprimer PCR of Whole Plasmid. *Biotechniques* **2002**, *33*, 1033-1038.

(2) O'Maille, P.E.; Bakhtina, M.; Tsai, M. D. Structure-based Combinatorial Protein Engineering (SCOPE). *J. Mol. Biol.* **2002**, *4*, 677-691.

(3) Studier, F. W. Protein Production by Auto-Induction in High-Density Shaking Cultures. *Protein Expres. Purif.* **2005**, *41*, 207-234.

(4) Davisson, V. J.; Woodside, A. B.; Neal, T. R.; Stremler, K. E.; Muehlbacher, M.; Poulter, C. D. Phosphorylation of Isoprenoid Alcohols. *J. Org. Chem.* **1986**. *51*, 4768-4779.

(5) Carter, S. G.; Karl, D. W. Inorganic Phosphate Assay with Malachite Green: An Improvement and Evaluation. *J. Biochem. Bioph. Meth.* **1982**. *7*. 7-13.

(6) Thulasiram, H. V.; Erickson, H. K.; Poulter, C. D. Chimeras of Two Isoprenoid Synthases Catalyze All Four Coupling Reactions in Isoprenoid Biosynthesis. *Science* **2007**, *316*, 73-76.

(7) Chen, M.; Poulter, C. D. Characterization of Thermophilic Archaeal Isopentenyl Phosphate Kinases. *Biochemistry*. **2010**. *49*. 207-217

(8) Woodside, A. B.; Huang, Z.; Poulter, C. D. Trisammonium Geranyl Diphosphate. *Org. Synth.* **2003**. 211-219

(9) Hosfield, D. J.; Zhang, Y.; Dougan, D. R.; Brooun, A.; Tari, L. W.; Swanson, R. V.; Finn, J. Structural Basis for Bisphosphonate-Mediated Inhibition of Isoprenoid Biosynthesis. *J. Biol. Chem.* **2004**. *279*. 8526-8529.

(10) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera – A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**. *13*, 1605-1612.

(11) Sali, A.; Blundell, T. L. Comparative Protein Modeling By Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993**. *3*. 779-815.

(12) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graph. Model.* **2006**. *2*. 247-260.

(13) Dunbrack, R. L. Rotamer Libraries in the 21st Century. *Curr. Opin. Struct. Biol.* **2002**. *4*. 431-440.

(14) Wang, J.; Wang, W.; Kollman, P. A.; Caldwell, J. W.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**. *9*. 1157-1174.

CHAPTER 5

CONCLUSION

5.1 Introduction

The isoprenoid biosynthetic pathway is one of nature's largest and most diverse. With over 62,000 natural products, terpenes include many well-known groups such as sterols, ubiquinones, carotenoids, and essential oils.[1-3] Despite responsibility for a wide structural array of chemicals, isoprenoids are simply built by coupling 3-methyl-1-butyl (isoprene) units. Dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP) are the universal isoprene units coupled in one of four ways – chain elongation, branching, cyclopropanation, and cyclobutanation – to make the terpenoid skeletal backbones.[2] Chain elongation is considered a regular or head-to-tail coupling, while the others are irregular or nonhead-to-tail.

The type I isoprenoid synthase (IS-1) fold is intimately involved in the isoprenoid biosynthetic pathway. The all α-helical IS-1 fold has been observed in multiple *E*-polyprenyl diphosphate synthases, enzymes catalyzing nonhead-to-tail reactions, and terpenoid cyclases, making enzymes exhibiting the structure members of a superfamily.[2,4-7] The prevalence of the IS-1 fold suggests divergent evolution from a common primordial ancestor early in the evolution of terpene biosynthesis. Five conserved regions of *E*-chain elongation enzymes exist and form the active sites of the

enzymes with few exceptions.[8]

Genes for the closely related farnesyl and chrysanthemyl diphosphate synthases (FPPase and CPPase, respectively) were isolated from snowfield sagebrush *Artemisia tridentata* ssp. spiciformis, both enzymes being IS-1 fold homodimers.[9] FPPase forms the regularly coupled, C_{15} compound farnesyl diphosphate with high efficiency and specificity, whereas CPPase is relatively inefficient and catalytically promiscuous exhibiting all the fundamental couplings in its products (geranyl diphosphate, chain elongation; chrysanthemyl diphosphate, cyclopropanation; lavandulyl diphosphate, branching; maconelliyl diphosphate, cyclobutanation). Despite an apparent catalytic gap between the enzymes, their sequence alignments show a striking 69% identity and 84% similarity, not only implying a recent divergence of the genes from a common origin, but also perhaps the evolution of cyclopropanation activity from an ancestral *E*-chain elongation enzyme.[8] The staggering sequence and presumed structural resemblance, yet contrasting catalytic activities between the sagebrush enzymes yields a unique opportunity to probe structure-function aspects of the IS-1 fold.

The use of chimeric proteins has proven utility in the enzymologist's toolbox. In the case of FPPase and CPPase chimeragenesis, the tool is used to mimic evolution using DNA sequences that underwent the scrutiny of natural selection. Sagebrush FPPase and CPPase chimeras were built around the α-helices and loops of the IS-1 fold in an attempt to deduce the structural elements in the IS-1 fold responsible for catalytic selectivity and activity. Chimeras were assembled in both directions (turning FPPase into CPPase, and vice-versa) and done so using PCR in order to produce constructs that maintain wild-type enzymatic sequences. Product ratio and kinetic analyses were

performed, if possible, in addition to molecular modeling of the parental enzymes and select chimeric enzyme constructs.

5.2 The N- to C-terminal Conversion of FPPase to CPPase

Four milestones along the N- to C-terminal transformation of FPPase to CPPase occur. First, approximately 19-, 5-, and 11-fold increases of $K_M^{DMAPP,EA}$, $K_M^{IPP,ED}$, and $K_M^{GPP,EA}$, respectively, were observed at the C1 to C2 chimeric enzyme transition, associated with an N53V FPPase to CPPase mutation in the first conserved region of E-chain elongation enzymes (EA, binding to the electron acceptor site; ED, binding to the electron donor site; GPP, geranyl diphosphate). The second milestone is marked by chain elongation activity transitioning from preferential farnesyl diphosphate (C_{15}) formation to geranyl diphosphate (C_{10}) at chimera C10 due to an over 100-fold increase of $K_M^{GPP,EA}$. The loss of GPP binding ability observed with C10 appears to be the result of multiple mutations involving the helices that form the GPP binding pocket and a T171R FPPase to CPPase active-site mutation.

The third milestone appears at the C10 to C11 transition, where an approximately 10-fold decrease in $k_{cat}^{DMAPP/IPP}$, a 17-fold increase in $K_M^{DMAPP, EA}$, and a 22-fold increase in $K_M^{IPP,ED}$ occur, dropping the catalytic efficiency to CPPase-like levels. The drastic reduction in catalytic efficiency (k_{cat}/K_M) at the C10 to C11 chimeric enzyme transition is associated with a T194G active-site FPPase to CPPase mutation in the fourth *E*-chain elongase conserved region. The fourth and final milestone in the FPPase to CPPase metamorphosis transpires as significant irregular coupling ability appears at chimera C13, indicating a CPPase-like allowance of DMAPP binding to the electron donor site. The

introduction of irregular coupling ability correlated with F231Y and D235N FPPase to CPPase mutations in the fifth conserved region. In contrast to CPPase, C13 preferentially formed the branched product lavandulyl diphosphate (LPP) over the cyclopropyl product chrysanthemyl diphosphate (CPP), indicating the C-terminal region outside the IS-1 fold is a significant determinant in the catalytic selectivity of the enzymes.

5.3 The N- to C-terminal Conversion of CPPase to FPPase

Along the N- to C-terminal conversion of CPPase to FPPase, selective binding of IPP to the electron donor site, and thus catalytic propensity towards the chain elongation reaction, was regained at the F1 chimera, in addition to a recovery in GPP binding ability. Similarly to CPPase, the F1 chimera preferentially forms the cyclopropyl product CPP over the branched product LPP when condensing two DMAPP molecules. Significant irregular coupling ability was lost at chimera F2. In this, only chimeras containing complete IS-1 fold sequence from CPPase did not selectively bind IPP to the electron donor site, and therefore competitively catalyzed regular and irregular couplings when incubated with DMAPP and IPP. An FPPase-like catalytic efficiency was not restored at any point in the CPPase and FPPase transition, indicating the significance of the C-terminal region in defining catalytic activity.

5.4 The Significance of the C-terminal Region in Defining the Catalytic Selectivity and Activity of CPPase and FPPase

Chimera C13 preferentially formed LPP when irregularly condensing two DMAPP molecules, whereas CPPase preferentially formed CPP. C13 differs from

CPPase in that the C-terminal region contains sequence from FPPase. In addition, swapping the C-terminal region of FPPase with CPPase sequence (to produce chimera F13) reduced the catalytic efficiency over 8500-fold to CPPase-like levels. The C-terminal regions of FPPase and CPPase differ by two active site residues – D256E and R343G. Although the mutations may not bear sole responsibility for the affect of the C-terminal region, D256E and R343G play substantial roles in the active site, and in turn, likely affect the catalytic activity and selectivity of the sagebrush enzymes.

5.5 References

(1) Dictionary of Natural Products. http://dnp.chemnetbase.com/dictionary-search/results.do?id=7816252&props=&struct=start&disp=&si (accessed 01/08/15).

(2) Thulasiram, H. V.; Erickson, H. K.; Poulter, C. D. A Common Mechanism for Branching, Cyclopropanation, and Cyclobutanation Reactions in the Isoprenoid Biosynthetic Pathway. *J. Am. Chem. Soc.* **2008**, *130,* 1966-1971.

(3) Tarshis, L. C.; Proteau, P. J.; Kellogg, B. A.; Sacchettini, J. C.; Poulter, C. D. Regulation of Product Chain Length by Isoprenyl Diphosphate Synthases. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 15018-15023.

(4) Poulter, C. D. Farnesyl diphosphate synthase. A paradigm for understanding structure and function relationships in E-polyprenyl diphosphate synthases. *Phytochem. Rev.* **2006**, *5*, 17-26.

(5) Pandit, J.; Danley, D. E.; Schulte, G. K.; Mazzalupo, S.; Pauly, T. A.; Hayward, C. M.; Hamanaka, E. S.; Thompson, J. F.; Harwood, Jr., H. J. Crystal Structure of Human Squalene Synthase. *J. Biol. Chem.* **2000**, *275*, 30610-30617.

(6) Rivera, S. B.; Swedlund, B. D.; King, G. J.; Bell, R. N.; Hussey, Jr., C. E.; Shattuck-Eidens, D. M.; Wrobel, W. M.; Peiser, G. D.; Poulter, C. D. Chrysanthemyl Diphosphate Synthase: Isolation of the gene and characterization of the recombinant non-head-to-tail monoterpene synthase from *Chrysanthemum cinerariaefolium*. *Proc. Natl. Acad. Sci.* **2001**, *98*, 4373-4378.

(7) Christianson, D. W. Structural Biology and Chemistry of the Terpenoid Cyclases. *Chem. Rev.* **2006**, *106*, 3412-3442

(8) Thulasiram, H. V.; Erickson, H. K.; Poulter, C. D. Chimeras of Two Isoprenoid Synthases Catalyze All Four Coupling Reactions in Isoprenoid Biosynthesis. *Science* **2007**, *316*, 73-76.

(9) Hemmerlin, A.; Rivera, S. B.; Erickson, H. K.; Poulter, C. D. Enzymes Encoded by the Farnesyl Diphosphate Synthase Gene Family in the Big Sagebrush *Artimisia tridentata* ssp. spiciformis. *J. Biol. Chem.* **2003**, *278*, 32132-32140.

APPENDIX

SUPPLEMENTARY INFORMATION, FIGURES, AND TABLES

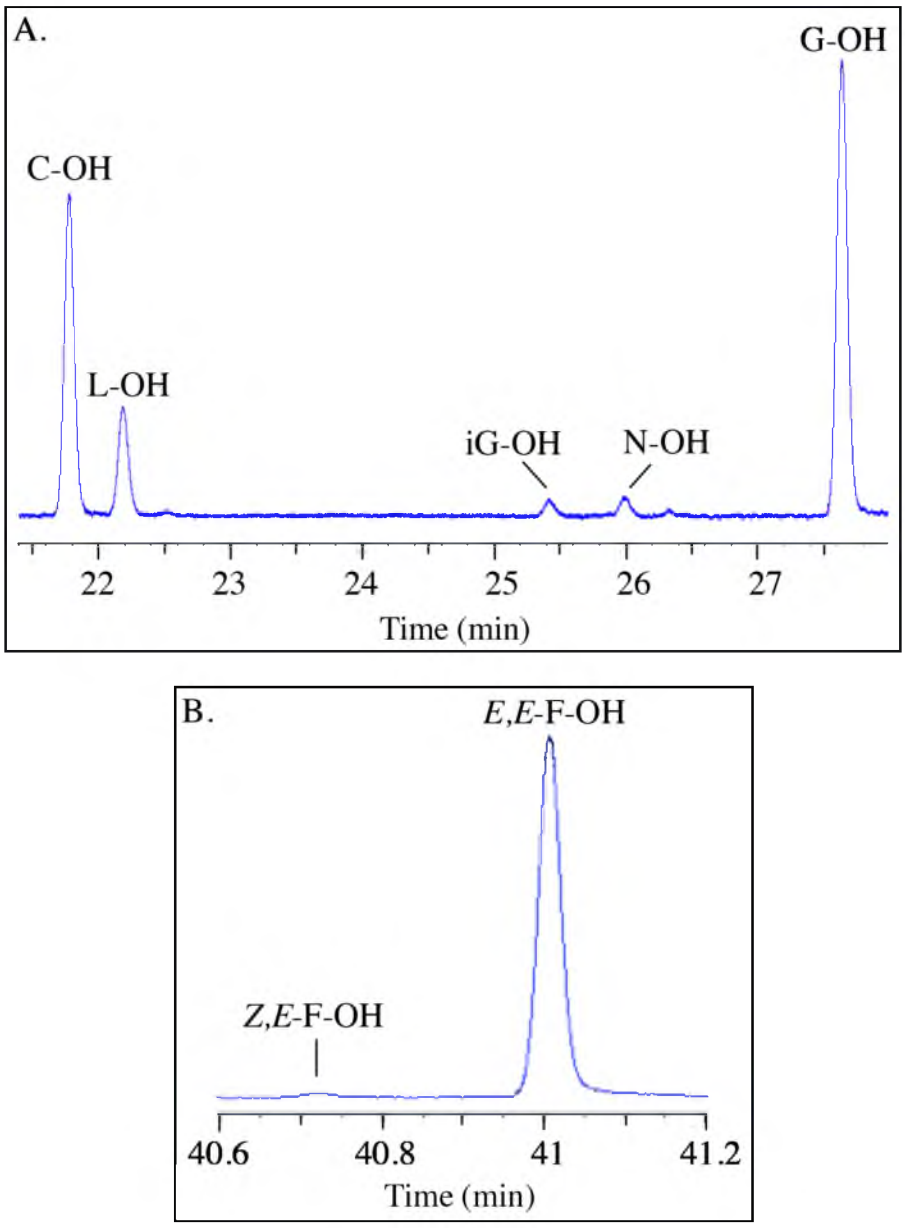


Figure A1. Representative GC-FID traces of regular and irregular coupling, alcohol products (post-calf intestinal phosphatase digestion). **A.** C_{10} products of CPPase under incubation with 500 μM of DMAPP and IPP (C-OH, chrysanthemol, ~ 21.8 min; lavandulol, L-OH, ~ 22.2 min; iG-OH, 3-methylene geraniol, ~ 25.4 min; N-OH, nerol, ~ 25 min; G-OH, geraniol, ~ 27.7 min; Agilent HP-5 column). **B.** C_{15} products of FPPase under incubation with 500 μM DMAPP and IPP (*Z,E*-F-OH, *Z,E*-farnesol, ~ 40.7 min; *E,E*-F-OH, *E,E*-farnesol, ~ 41 min; Agilent HP-5 column).

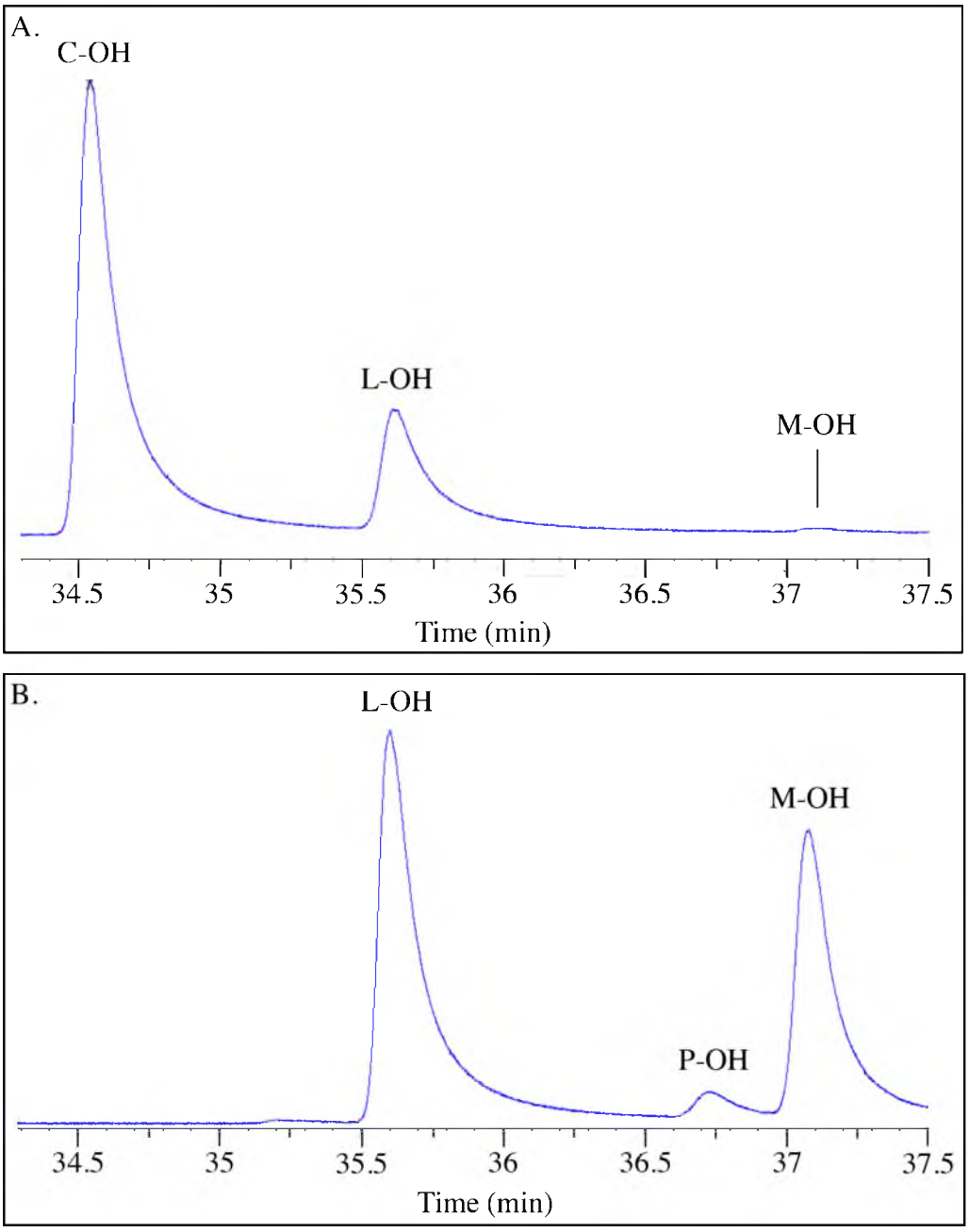


Figure A2. Representative GC-FID traces of irregularly coupled products (post-calf intestinal phosphatase digestion). **A, B.** Irregular coupling products from incubation of CPPase and chimera C6, respectively, with 3 mM of DMAPP (C-OH, chrysanthemol, ~34.5 min; L-OH, lavandulol, ~35.6 min; P-OH, planococcol, ~36.7 min; M-OH, maconelliol, ~37.1 min; Agilent HP Chiral column).

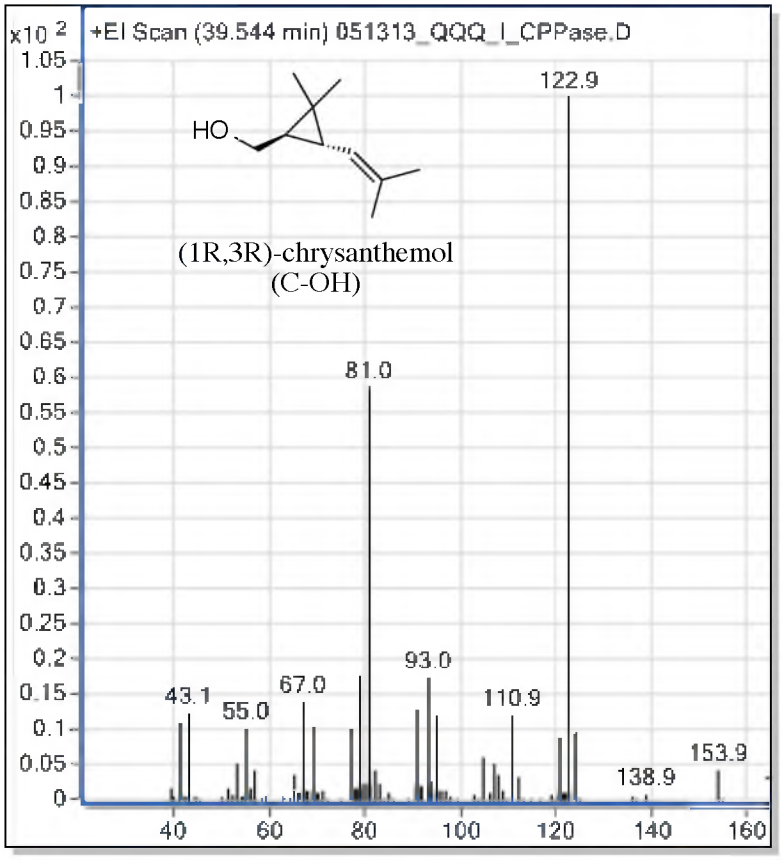


Figure A3. Electron impact fragmentation of chrysanthemol. Fragmentation pattern obtained from GC-MS analysis of irregular product formation by CPPase.

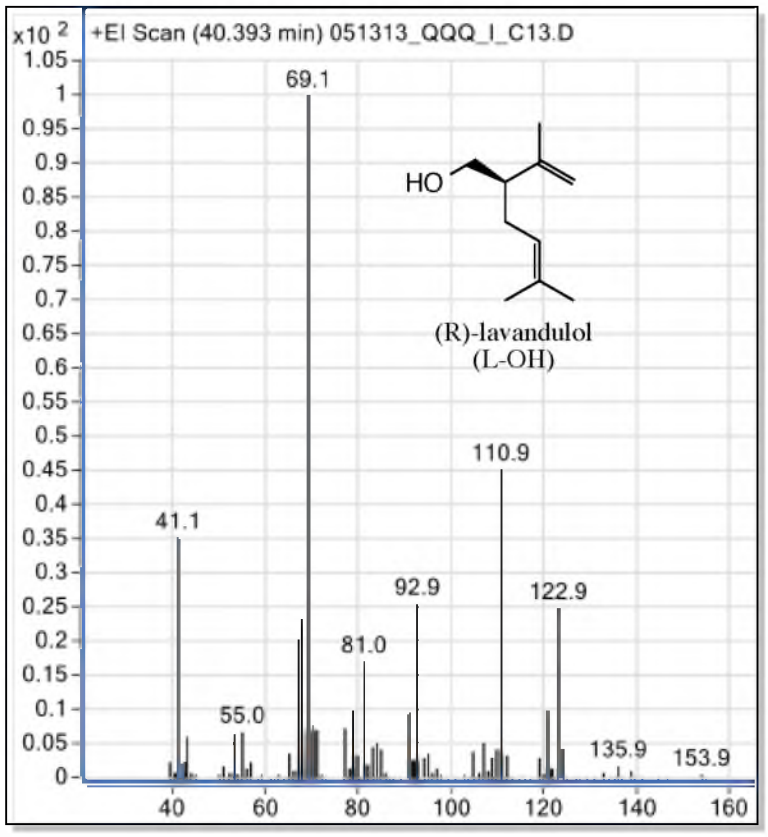


Figure A4. Electron impact fragmentation of lavandulol. Fragmentation pattern obtained from GC-MS analysis of irregular product formation by chimera C13.

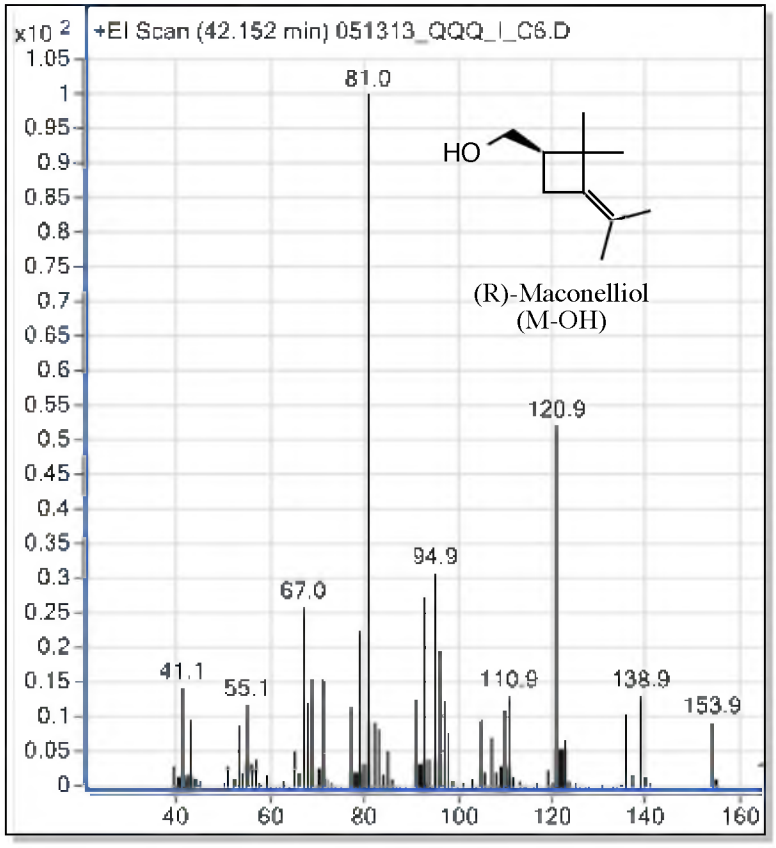


Figure A5. Electron impact fragmentation of maconelliol. Fragmentation pattern obtained from GC-MS analysis of irregular product formation by chimera C6.

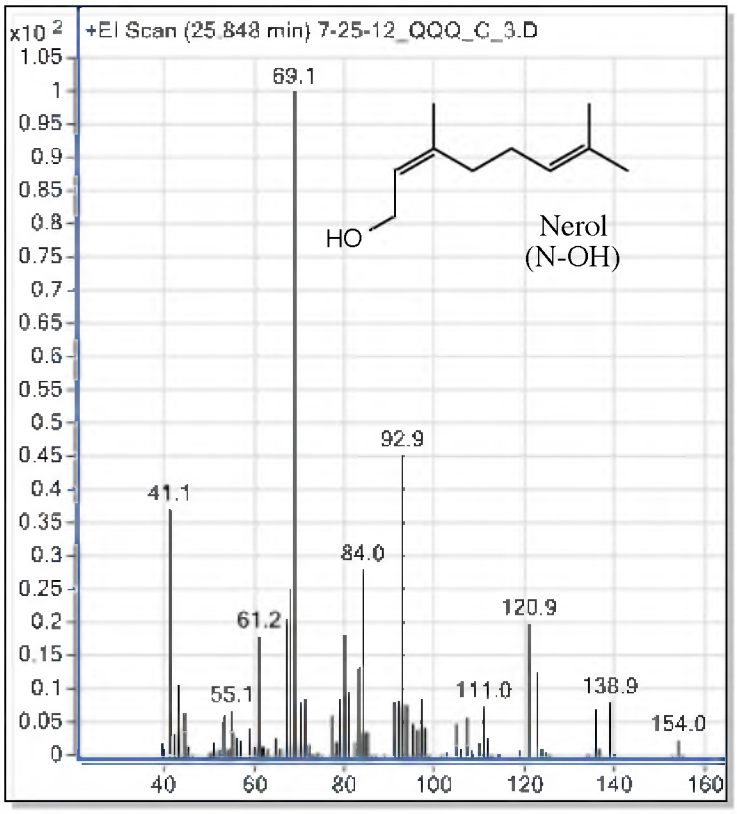


Figure A6. Electron impact fragmentation of nerol. Fragmentation pattern obtained from GC-MS analysis of regular product formation by CPPase.

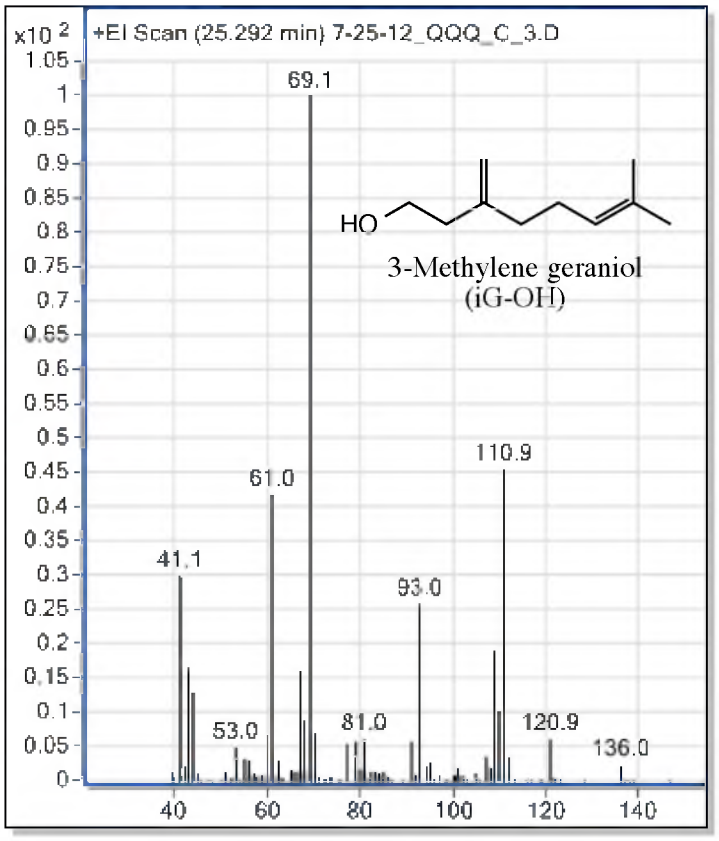


Figure A7. Electron impact fragmentation of 3-methylene geraniol. Fragmentation pattern obtained from GC-MS analysis of regular product formation by CPPase.

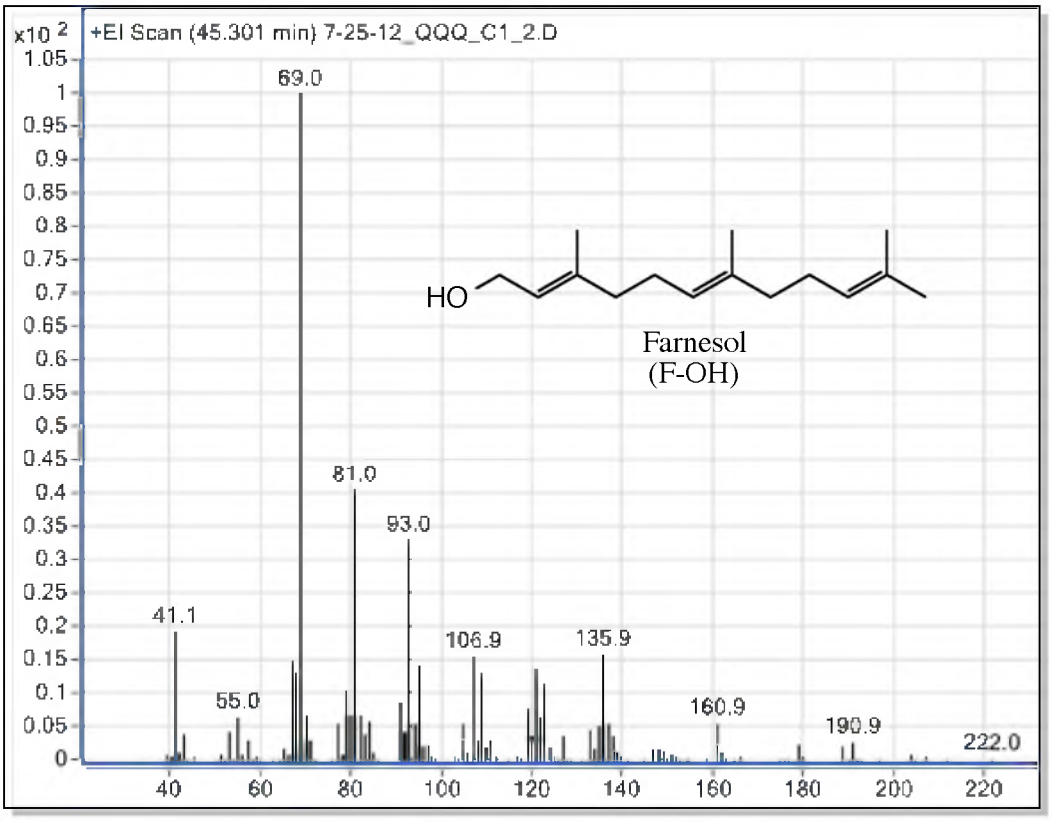


Figure A8. Electron impact fragmentation of farnesol. Fragmentation pattern obtained from GC-MS analysis of regular product formation of chimera C1.

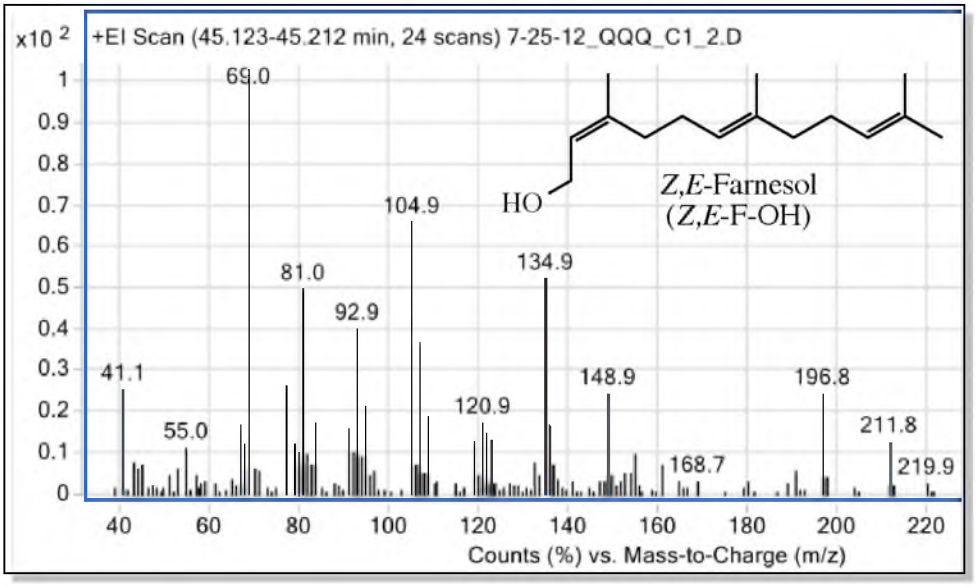


Figure A9. Electron impact fragmentation of *Z,E*-farnesol. Fragmentation pattern obtained from GC-MS analysis of regular product formation of chimera C1.

CPPase:
5′-CATATGACGACGACGCTGTCTAGTAATCTGAACTCGCAATTTATGCAAGTCTACGA
AACCCTGAAATCTGAACTGATCCACGACCCGCTGTTTGAATTTGATGACGATAGTCGTC
AGTGGGTGGAACGCATGATTGACTATACCGTTCCGGGCGGTAAAATGGTCCGTGGCTAT
AGTGTGGTTGATTCCTACCAACTGCTGAAAGGTGAAGAACTGACCGAAGAAGAAGCGTT
TCTGGCGTGCGCCCTGGGCTGGTGTACGGAATGGTTTCAGGCTTTCATCCTGCTGCATG
ACGATATGATGGATGGCTCTCACACGCGTCGCGGTCAACCGTGCTGGTTTCGTCTGCCG
GAAGTTGGCGCAGTCGCTATTAACGATGGTGTGCTGCTGCGTAATCATGTTCACCGCAT
CCTGAAAAAACATTTCCAGGGTAAAGCGTATTACGTGCACCTGGTTGACCTGTTTAACG
AAACCGAATTCCAGACGATTTCAGGCCAAATGATTGATACCATCTCGCGCCTGGCCGGT
CAGAAAGAACTGAGTAAATATAGCATGTCTCTGAACCGTCGCATTGTTCAATATAAAGG
CGCATATTACTCCTGCTACCTGCCGATCGCGTGTGCCCTGCTGATGTTTGGTGAAAATC
TGGACGATTATGTCCAGGTGAAAGATATTCTGGTCGAACTGGGCATGTATTACCAGATC
CAAAACGACTACCTGGATACCTTTGGTGACCCGAATGTTTTCGGCAAAACCGGTACGGA
TATTGAAGAATGCAAATGTTCATGGCTGATCGCAAAAGCTCTGGAACTGGCCAACGAAG
AACAGAAGAAAATTCTGTCGGAAAATTATGGCATCAAAGACCCGGCGAAAGTCGCCAAA
GTGAAAGAAATTTACCATGCGCTGAATCTGAAAGGTGCCTATGAAGATTACGAAACGAA
CCTGTATGAAAATAGCATGAAAGCAATCAAAGCTCACCCGTCTATTGCGGTGCAGGCAG
TGCTGAAATCTTGTCTGGAAAAAATGTATAAAGGTCATAAATAAGGATCC-3′

Figure A10. DNA sequence of CPPase synthesized in *E. coli* context.

FPPase:
5′-CATATGAGCAGCAGCAAAAGCATTGACCTGAAAAGCAAATTCCTGAAAGTTTACGA
CACCCTGAAAAGCGACCTGATTAACGACCCGGCATTCGAATTTGATGACGATAGCCGTC
AGTGGATTCAAAAAATGCTGGACTATAACGTTCCGGGCGGTAAACTGAATCGCGGCCTG
TCAGTGGTTGATTCGTACCAGCTGCTGAAAGGCGGTGAACTGAGCGACGATGAAATTTT
TCTGAGCTCTGCACTGGGCTGGTGCATCGAATGGCTGCAGGCTTATTTCCTGGTGCTGG
ACGATATTATGGATGAATCTCATACCCGTCGCGGTCAACCGTGTTGGTTTCGTCTGCCG
AAAGTCGGCATGATTGCGGCCAACGATGGTATCCTGCTGCGTAATCATGTGCCGCGCAT
CCTGAAAAAACACTTCCGCGGCAAACCGTATTACGTCGACCTGGTGGATCTGTTTAACG
AAGTTGAATTCCAGACCGCGAGTGGCCAAATGATTGACCTGATCACCACCCTGGTGGGC
GAAAAAGATCTGTCCAAATATAGTCTGTCCATTCATCGTCGCATCGTTCAGTATAAAAC
GGCCTATTACTCATTTTACCTGCCGGTCGCATGCGCTCTGCTGATGTTCGGTGAAGACC
TGGATAAACACGTTGAAGTCAAAAATGTGCTGGTTGAAATGGGCACCTATTTTCAGGTG
CAAGACGATTACCTGGATTGCTTCGGTGCGCCGGAAGTTATTGGCAAAATCGGTACGGA
CATCGAAGATTTTAAATGTTCCTGGCTGGTCGTGAAAGCACTGGAACTGGCTAACGAAG
AACAGAAGAAAACCCTGCATGAAAACTACGGTAAAAAAGATCCGGCGTCAGTTGCCAAA
GTCAAAGAAGTGTACCACACGCTGAACCTGCAAGCGGTTTTCGAAGATTATGAAGCCAC
CTCGTACAAAAAACTGATTACGAGCATCGAAAATCACCCGTCTAAAGCGGTCCAAGCAG
TTCTGAAATCATTCCTGGGTAAAATCTACAAACGCCAAAAATAAGGATCC-3′

Figure A11. DNA sequence of FPPase synthesized in *E. coli* context.

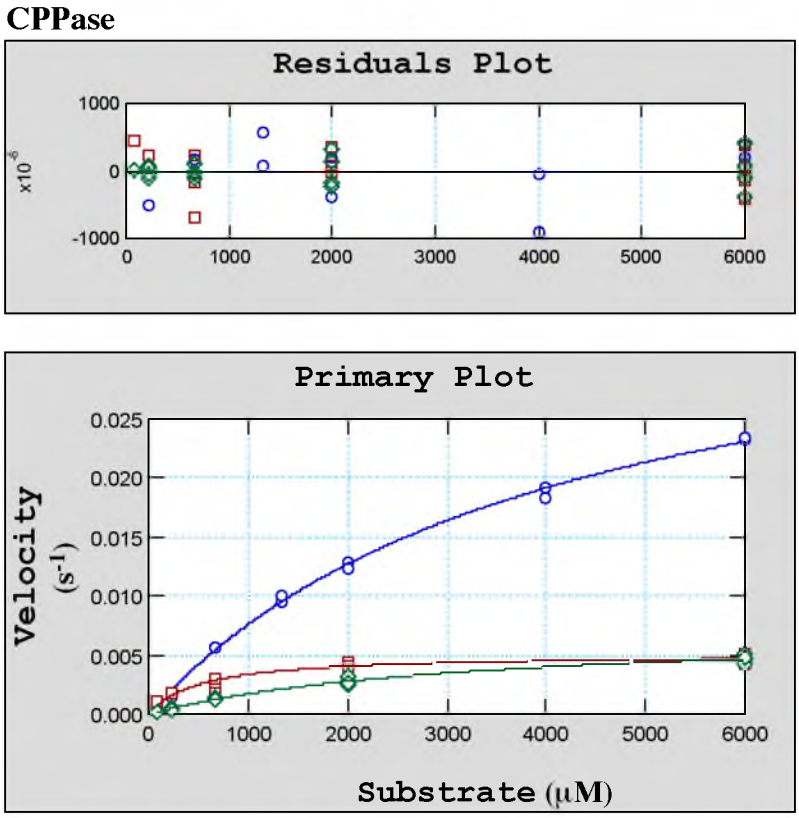


Figure A12. Hyperbolic plots for estimating the apparent Michaelis-Menten kinetic values of CPPase for the irregular coupling (varying DMAPP, blue) and chain elongation reactions (varying DMAPP at 6 mM IPP, brown; varying IPP at 6 mM DMAPP, green).

FPPase

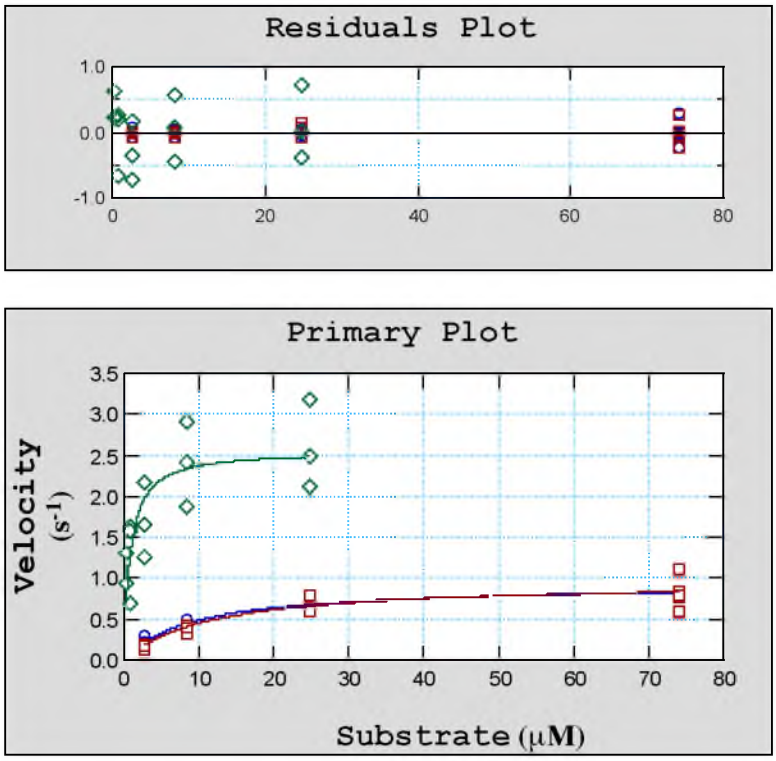


Figure A13. Hyperbolic plots for estimating the apparent Michaelis-Menten kinetic values of FPPase for the chain elongation reaction (varying GPP at 74 μM IPP, green; varying DMAPP at 74 μM IPP, blue; varying IPP at 74 μM DMAPP, brown).

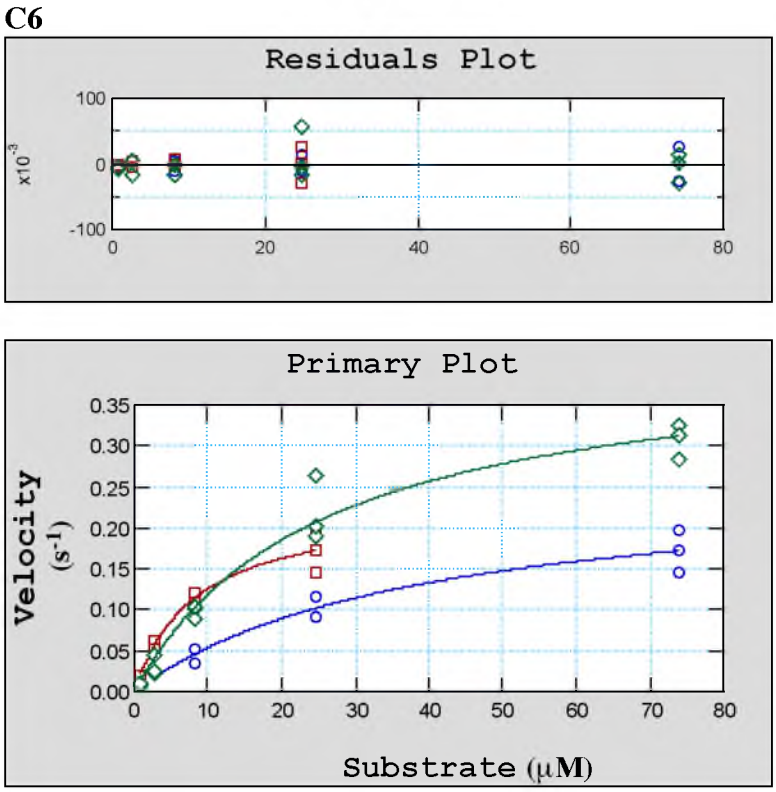


Figure A14. Hyperbolic plots for estimating the apparent Michaelis-Menten kinetic values of chimera C6 for the chain elongation reaction (varying GPP at 25 μM IPP, green; varying DMAPP at 25 μM IPP, blue; varying IPP at 74 μM DMAPP, brown).

F8

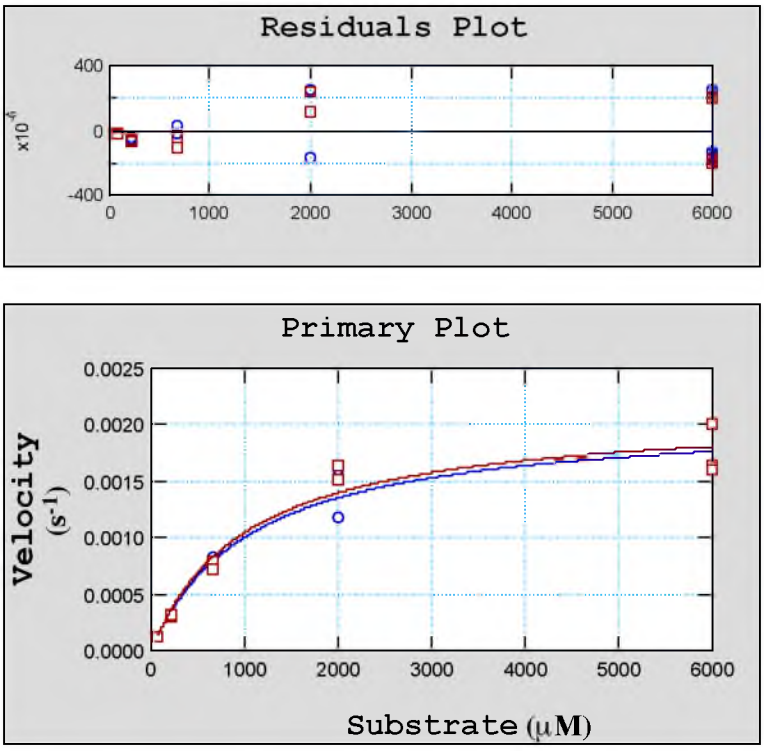


Figure A15. Hyperbolic plots for estimating the apparent Michaelis-Menten kinetic values of chimera F8 for the chain elongation reaction (varying DMAPP at 6 mM IPP, blue; varying IPP at 6 mM DMAPP, brown).

Table A1. Forward and reverse crossover primers used to form the megaprimers in the first step of the PCR chimeragenesis protocol. Reverse crossover primers are named as to reflect the chimeric enzyme they will produce, for example, the F_3_C primer was used to build the F3 chimera. (FP indicates forward primer).

PCR Primer	Sequence (5'-3')
C_1_F	CCACTGACAGGCCGCGATTCATTTTACCGCCCGGAACGGTATAG
F_1_C	GGAATCAACCACACTATAGCCACGGACCAGGCCGCGATTCAGTTTACCG
C_2_F	GTCGCTCAGTTCACCGCCTTTCAGCAGTTGGTAGGAATCAACCACA
F_2_C	CGCTTCTTCTTCGGTCAGTTCTTCACCTTTCAGCAGCTGGTACGAATCAACCA
C_3_F	CAGTGCAGAGCTCAGAAAAATTTCATCGTCGGTCAGTTCTTCACCTTTCAGCAGTTG
F_3_C	GCACGCCAGAAACGCTTCTTCTTCGCTCAGTTCACCGCCTTTCAGC
C_4_F	TGACCGCGACGGGTATGAGATTCATCCATCATATCGTCATGCAGCAGGATG
F_4_C	CGACGCGTGTGAGAGCCATCCATAATATCGTCCAGCACCAGGAAATAAG
C_5_F	GATACCATCGTTGGCCGCAATCATGCCAACTTCCGGCAGACG
F_5_C	GCACACCATCGTTAATAGCGACTGCGCCAACTTCCGGCAGACG
C_6_F	CGACGTAATACGGTTTGCCGCGGAAATGTTTTTTCAGGATGCGGTGAACATGATTAC
F_6_C	GGTGCACGTAATACGCTTTACCCTGGAAGTGTTTTTTCAGGATGCGCGG
C_7_F	ACAGATCCACCAGGTCGACGTAATACGCTTTACCCTGGAAATGTTTTTTCAGG
F_7_C	CAGGTCAACCAGGTGCACGTAATACGGTTTGCCGCGGAAGTG
C_8_F	CGCCCACCAGGGTGGTGATGGTATCAATCATTTGGCCTGAAATCGTC
F_8_C	CGGCCAGGCGCGAGATCAGGTCAATCATTTGGCCACTCG
C_9_F	GAACGATGCGACGATGAATGGACAGGCTATATTTACTCAGTTCTTTCTGACCGGCC
F_9_C	GAACAATGCGACGGTTCAGAGACATACTATATTTGGACAGATCTTTTTCGCCCACC
C_10_F	CGGCAGGTAAAATGAGTAATAGGCCGTTTTATATTGAACAATGCGACGGTTCAGAGACATG
F_10_C	CAGGTAGCAGGAGTAATATGCGCCTTTATACTGAACGATGCGACGATGAATGGAC
C_11_F	GACTTCAACGTGTTTATCCAGGTCTTCACCAAACATCAGCAGGGCACACGC
F_11_C	ACCTGGACATAATCGTCCAGATTTTCACCGAACATCAGCAGAGCGCATGC
C_12_F	CCCATTTCAACCAGCACATTTTTGACTTCAACATAATCGTCCAGATTTTCACCAAACATCAGCAG
F_12_C	CAGTTCGACCAGAATATCTTTCACCTGGACGTGTTTATCCAGGTCTTCACCGAACATCAG
C_13_F	CCAATAACTTCCGGCGCACCAAAGGTATCCAGGTAGTCGTTTTGGATCTG
F_13_C	TGCCGAAAACATTCGGGTCACCGAAGCAATCCAGGTAATCGTCTTGCAC
CPPase_FP	GGCCTGGTGCGCGCGGCAGCCATATGACGACGACGCTGTCTAGTAATCTG
FPPase_FP	GGCCTGGTGCGCGCGGCAGCCATATGAGCAGCAGCAAAAGCATTGACC

Table A2. Representative electrospray ionization mass spectrometry (ESI-MS) analyses for molecular weight verification of the enzymes.

Enzyme	Predicted MW (Da)	Observed MW (Da)
C1	41842.8	41841.6
C3	41963.9	41964
C4	41977	41978.4
C5	41906.8	41907.2
C6	41900.7	41901.6
C10	41921.7	41922.4
C11	41847.7	41848.8
C13	41912.7	41914.4
f_C_f	41795.7	41796.8
F9	41620.6	41619.2
F10	41623.2	41625.6
F11	41699.7	41699.2
F12	41687.7	41688
F13	41634.6	41633.6
c_F_c	41751.7	41750.4